DATE DUE

~~MAR 2 - 1993~~		~~AUG 1 9 1991~~	
~~DEC 1 0 1989~~			
retd	12/19/89	~~FEB 2 7 2001~~	
~~JAN 3 1 1990~~			
~~APR 3 0 1990~~			
~~NOV 3 0 1990~~			
~~DEC 2 7 1990~~			
~~MAY 3 1 1991~~			
~~DEC 0 7 1991~~			
GAYLORD			PRINTED IN U.S.A.

CORNELL UNIVERSITY

Department of
Theoretical + Applied Mechanics
The College of Engineering
Thurston Hall
CORNELL UNIVERSITY
Ithaca, N. Y. 14850

ENGINEERING MATHEMATICS

Volume 2

Ithaca, New York

H. D. Block
E. T. Cranch
P. J. Hilton
R. J. Wal

The preparation of this material has been partially supported by Grant NSF-GE2564 made to Cornell University by the National Science Foundation.

Acknowledgments

The authors wish to express their thanks to their technical assistants in the preparation of this second volume: to Mr. Thomas Kuffel for general assistance, to our printer Mr. Harry Gleason, and to Mrs. Anne Kramer, who did all the typing. Most of all we acknowledge our indebtedness to Mrs. Sandria Kerr for her ever-cheerful handling of the editing, illustrating, and proof-reading of both volumes.

H. D. Block

E. T. Cranch

P. J. Hilton

R. J. Walker

Ithaca, New York

January 1965

CHAPTER 5

Vector Spaces With an Inner Product

Table of Contents

1. Physical Vectors in Three-Dimensional Space 5.1
2. Inner Product in Abstract Vector Spaces 5.6
3. Orthogonal Bases. 5.14
4. Matrices, An Application. 5.33
5. Eigenvectors and Eigenvalues. 5.44
6. Expansion in Eigenvectors. Numerical Computation 5.68
7. Change of Basis 5.80
8. Orthogonal Transformations. 5.86
9. Quadratic Forms 5.98
10. Vector Spaces Over the Complex Field. 5.111
11. Self-Adjoint Operators. 5.116

CHAPTER 6

Linear Differential Equations with Constant Coefficients

Table of Contents

1. Two First Order Homogeneous Simultaneous
Differential Equations. 6.1
2. Systems of First Order Homogeneous Linear
Differential Equations with Constant
Coefficients. General Case 6.17
3. Two First Order Nonhomogeneous Simultaneous
Differential Equations. 6.47
4. Systems of First Order Nonhomogeneous
Equations. General Case. 6.76
5. Systems of Second Order Differential
Equations 6.101
6. A Single Linear Differential Equation
with Constant Coefficients. 6.126

CHAPTER 7

Vector Analysis in 3-Space

Table of Contents

1. Vector Algebra in 3-Space. 7.1
2. Derivatives and Partial Derivatives. 7.18
3. Scalar and Vector Fields and the
   Differential Operator ∇. 7.38
4. Algebra of the Operator ∇. 7.65
5. Charge of Coordinates. 7.70

CHAPTER 8

Boundary Value Problems and Fourier Series

Table of Contents

1. Steady State Heat Distribution in a
   Rectangular Plate 8.1
2. Fourier Series on the Interval $(-\pi,\pi)$ 8.17
3. Change of Interval. 8.37
4. Fourier Sine and Cosine Series. 8.40
5. Applications to Solution of Differential
   Equations 8.46

CHAPTER 5

Vector Spaces With An Inner Product

In Chapter 2 we defined a vector space in terms of two operations, addition of two vectors and multiplication of a vector by a real number (or scalar). In terms of these operations an algebra of vectors was developed, involving such concepts as linear dependence, linear subspaces, and dimension. Linear transformations were defined as transformations which preserved the two basic operations. Finally, matrices were introduced as representing linear transformations in finite dimensional spaces, and from this matrix algebra was deduced.

Throughout all this development no mention was made of the length of a vector, or of the angle between two vectors. The time has now come for us to incorporate these concepts into our theory. This we can do most easily by introducing a new operation, the inner product, with appropriate specified properties.

As in the earlier chapter we start out with a physical motivation based on geometrical properties of vectors in two or three dimensions.

1. Physical Vectors in Three-Dimensional Space

As in Sections 2 and 4 of Chapter 2 we first consider vectors as forces through a point, displacements, or other physical quantities, representing them by arrows from the origin. If we settle on a unit of measurement (pound, foot, etc.) each such vector has associated with it a non-negative number called its magnitude or its <u>length</u>. Following the most modern notation we denote the length of the vector $\vec{v}$ by $\|\vec{v}\|$. (Some writers use $|\vec{v}|$ but this can cause confusion with absolute value or determinant.) The basic properties of length are

(a) $\|a\vec{v}\| = |a| \ \|\vec{v}\|$ for any scalar a,

(b) $\|\vec{v}\| \begin{cases} = 0 & \text{if } \vec{v} = \vec{0}, \\ > 0 & \text{if } \vec{v} \neq \vec{0}, \end{cases}$

(c) $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$.

Property (c) is known as the "triangle inequality" and follows from the parallelogram law of vector addition (see Figure 1.1).

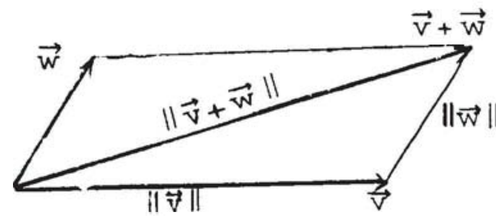


Figure 1.1

Between any two <u>non-zero</u> vectors, $\vec{v}$ and $\vec{w}$, there is a unique angle θ in the range $0 \leq \theta \leq \pi$ (Figure 1.2).

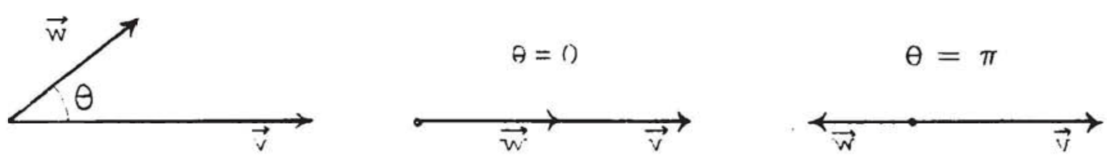


Figure 1.2

In terms of this angle and the lengths of the vectors we define a new algebraic operation as follows:

$$\vec{v}\cdot\vec{w} = \|\vec{v}\|\,\|\vec{w}\|\cos\theta$$

This is called the dot or scalar or inner product of the two vectors; we shall generally use the last of these three names. In order to remove the restriction that $\vec{v} \neq \vec{0}$ and $\vec{w} \neq \vec{0}$ we further define

$$\vec{v}\cdot\vec{0} = \vec{0}\cdot\vec{w} = 0$$

for any vectors $\vec{v}$ and $\vec{w}$.

The inner product has the following properties:

(i) $\vec{v}\cdot\vec{w}$ is a scalar,

(ii) $\vec{v}\cdot\vec{w} = \vec{w}\cdot\vec{v}$,

(iii) $(a\vec{v})\cdot\vec{w} = \vec{v}\cdot(a\vec{w}) = a(\vec{v}\cdot\vec{w})$ for any scalar a,

(iv) $\vec{u}\cdot(\vec{v} + \vec{w}) = \vec{u}\cdot\vec{v} + \vec{u}\cdot\vec{w}$,

(v) $\vec{v}\cdot\vec{v} > 0$ if $\vec{v} \neq \vec{0}$.

(i), (ii) and (v) are obvious from the definition. We verify three cases of (iii):

If $a = 0$ then evidently each of the three expressions is zero;

If $a > 0$ then θ is the same in each of the three inner products and the relations obviously hold;

If $a < 0$ then θ is changed to its supplement in the first two products, thereby changing the sign of $\cos\theta$ and making the equalities come out correctly, in the light of (a).

There remains to prove the distributive law (iv). To do this note first (Figure 1.3) that

$$\vec{u}\cdot\vec{v} = \|u\|\ \|\vec{v}\|\ \cos\theta$$
$$= \epsilon\ \|\vec{u}\|\ \|\vec{v}'\|$$

where $\vec{v}'$ is the orthogonal projection of $\vec{v}$ on $\vec{u}$ and

$$\epsilon = \begin{cases} 1 \text{ if } 0 \leqq \theta \leqq \pi/2 \\ -1 \text{ if } \pi/2 < \theta \leqq \pi. \end{cases}$$

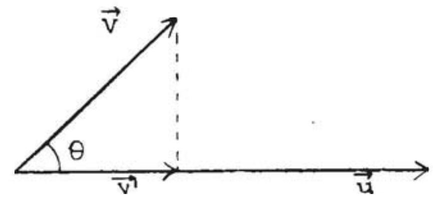


Figure 1.3

Now in Figure 1.4 a little geometry shows that $(\vec{v} + \vec{w})' = \vec{v}' + \vec{w}'$ in all cases, and from this the desired result is easily obtained (there are several cases to consider).

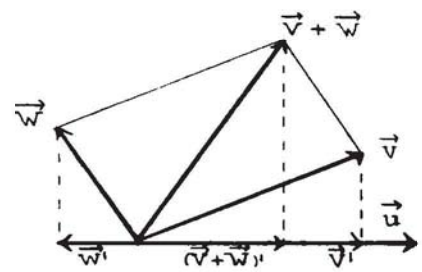


Figure 1.4

Example 1.1. A block sliding down an inclined plane (Figure 1.5) moves a distance $\vec{S}$ under the action of its weight $\vec{F}$. The work done by $\vec{F}$ in such a case is the product of the distance moved by the component of the force in the direction of motion. That is

$$\text{Work} = (\|\vec{F}\|\ \cos\theta)\ \|\vec{S}\|$$
$$= \vec{F}\cdot\vec{S}.$$

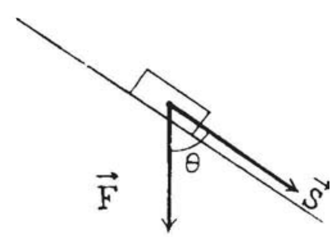


Figure 1.5

Example 1.2. In Example 8.3 of Chapter 4 the relation between the input $E(t) = E_0 e^{i\omega t}$ of the circuit

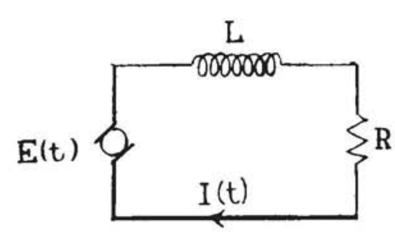


Figure 1.6

shown in Figure 1.6 and the response I(t) was shown to be given by Figure 1.7, where

$$I(t) = I_0 e^{i(\omega t - \theta)}$$

with

$$I_0 = E_0 \Big/ \sqrt{R^2 + L^2\omega^2}, \ \tan\theta = L\omega/R.$$

(See equation (8.32) of Chapter 4.) The power dissipation is given by $P = E \cdot I$, where E and I are considered as vectors. In this case

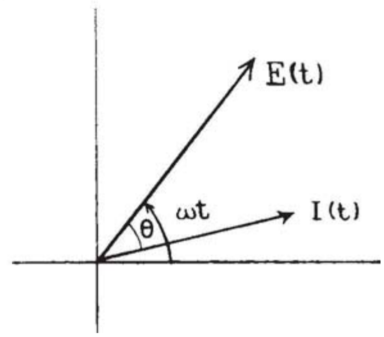


Figure 1.7

$$P = \|E\| \, \|I\| \cos\theta = E_0 I_0 \cos\theta$$

is constant, even though E and I vary with time. In particular,

$$P = E_0 \frac{E_0}{\sqrt{R^2+L^2\omega^2}} \ \frac{R}{\sqrt{R^2+L^2\omega^2}} = \frac{E_0{}^2 R}{R^2+L^2\omega^2} .$$

Let $\vec{i}$, $\vec{j}$, $\vec{k}$ be unit vectors along mutually perpendicular axes. Then any vector $\vec{v}$ can be expressed in terms of the basis $\vec{i}$, $\vec{j}$, $\vec{k}$ as

$$\vec{v} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k} .$$

If similarly,

$$\vec{w} = b_1\vec{i} + b_2\vec{j} + b_3\vec{k},$$

then

$$\begin{aligned} \vec{v}\cdot\vec{w} = {} & a_1b_1\vec{i}\cdot\vec{i} + a_1b_2\vec{i}\cdot\vec{j} + a_1b_3\vec{i}\cdot\vec{k} \\ & + a_2b_1\vec{j}\cdot\vec{i} + a_2b_2\vec{j}\cdot\vec{j} + a_2b_3\vec{j}\cdot\vec{k} \\ & + a_3b_1\vec{k}\cdot\vec{i} + a_3b_2\vec{k}\cdot\vec{j} + a_3b_3\vec{k}\cdot\vec{k}. \end{aligned} \tag{1.1}$$

Now $\vec{i}\cdot\vec{j} = \vec{i}\cdot\vec{k} = \vec{j}\cdot\vec{k} = 0$ since the vectors are perpendicular, and $\vec{i}\cdot\vec{i} = \vec{j}\cdot\vec{j} = \vec{k}\cdot\vec{k} = 1$ since they are unit vectors. Hence (1.1) reduces to

$$\vec{v}\cdot\vec{w} = a_1b_1 + a_2b_2 + a_3b_3 \ . \tag{1.2}$$

On the other hand, if the basis $\vec{i}$, $\vec{j}$, $\vec{k}$ were not of this special type the expression (1.1) could not in general be simplified. It is evident that such special bases, known as orthonormal bases, can be of great convenience in the algebra of inner products.

Problems

1.1 In 3-dimensional Euclidean space let (x_1,y_1,z_1) be the rectangular cartesian coordinates of point P_1 and (x_2,y_2,z_2) the coordinates of point P_2. If 0 is the origin of coordinates show that

$$\cos P_1OP_2 = \frac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{\left(x_1^{\,2} + y_1^{\,2} + z_1^{\,2}\right)\left(x_2^{\,2} + y_2^{\,2} + z_2^{\,2}\right)}} \ .$$

[Hint: Consider the vectors $\vec{OP}_1$ and $\vec{OP}_2$ and use equation (1.2).]

2. Inner Product in Abstract Vector Spaces.

In the previous section we defined the inner product of two vectors in Euclidean space in terms of geometric properties, distance and angle. If we wish to speak of an inner product of abstract vectors we cannot use this approach since we have no such geometric properties in a general abstract vector space. We therefore give an abstract definition, postulating the properties (properties (i),...,(v) of Section 1) that we wish our inner product to have. Then in terms of this inner product we will define length and angle.

Definition 2.1. An inner product in a vector space V is an operation $u \cdot v$ with the following properties:

1. If u and v are vectors of V then $u \cdot v$ is a scalar;
2. $u \cdot v = v \cdot u$;
3. For any scalar c, $(cu) \cdot v = u \cdot (cv) = c(u \cdot v)$;
4. $u \cdot (v + w) = u \cdot v + u \cdot w$;
5. $v \cdot v > 0$, if $v \neq 0$.

Example 2.1. Let V be the space V_n of n-tuples $(a_1, a_2, \ldots, a_n)$. Taking the clue from equation (1.2) we define

$$(2.1) \qquad (a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = a_1 b_1 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

The five properties are easily verified.

Example 2.2. Let V be the space C^0 of functions of x continuous on some interval $a \leqq x \leqq b$. If f and g are two such functions we define

$$(2.2) \qquad f \cdot g = \int_a^b f(x) g(x) dx.$$

Properties 1 to 4 are easily verified but 5 requires some consideration, as follows:

Since the 0 element of C^0 is the function which is identically zero, the condition $f \neq 0$ means that $f(x)$ is not identically zero on $a \leqq x \leqq b$. That is, there is some point c such that $f(c) \neq 0$. Hence $f(c)^2 > 0$, and since $f(x)$ is a continuous function there must be some interval $d \leqq x \leqq e$, containing c, over which $f(x)^2 > 0$. (Figure 2.1).

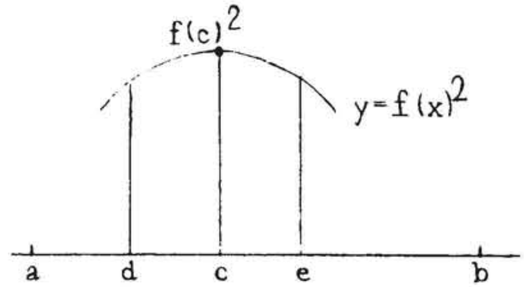


Figure 2.1

We have

$$f \cdot f = \int_a^b f(x)^2 dx = \int_a^d f(x)^2 dx + \int_d^e f(x)^2 dx + \int_e^b f(x)^2 dx.$$

The first and third integrals on the right hand side are non-negative, since we always have $f(x)^2 \geqq 0$, and from the choice of d and e the second integral is greater than zero. Hence $f \cdot f > 0$, proving property 5.

It should be noted that a given vector space may have more than one inner product defined in it. Only occasionally, however, does it become necessary to consider more than one inner product at a time. No such occasions will arise in this chapter, (but see Problem 2.14) and so in working with any vector space V we shall speak of <u>the</u> inner product of vectors in V. In particular, the inner products in V_n and in C^0 (or C^n, or C^∞) will always be the ones defined by equations (2.1) and (2.2).

Some further properties of inner products that follow readily from the definition are stated in the following theorem

Theorem 2.1. (a) For any vector $\vec{v}$, $\vec{0}\cdot\vec{v} = 0$.

(b) For any vectors w, $v_1,\ldots,v_m$, and scalars $a_1,\ldots,a_m$,

$$w\cdot \sum_{i=1}^{m} a_i v_i = \sum_{i=1}^{m} a_i (w\cdot v_i).$$

In terms of the inner product we can now define the notions of length and angle.

Definition 2.2. The length, or norm, of a vector v is $\|v\| = \sqrt{v\cdot v}$. The first two properties of length follow at once.

Theorem 2.2. $\|v\|$ has properties (a) and (b) of Section 1, i.e.,

(a) $\|av\| = |a|\,\|v\|$ for any scalar a,

(b) $\|v\| \begin{cases} = 0 & \text{if } v = 0, \\ > 0 & \text{if } v \neq 0. \end{cases}$

To prove the third property of length, the triangle inequality, we first obtain a preliminary result known as the Schwarz Inequality.

Theorem 2.3. $(v\cdot w)^2 \leqq (v\cdot v)(w\cdot w)$.

Proof. Consider the vector tv + w for any scalar t. We have

$$0 \leqq (tv + w)\cdot(tv + w) = (v\cdot v)t^2 + 2(v\cdot w)t + w\cdot w. \tag{2.3}$$

The right hand expression is a function of t which is never negative; hence its minimum value will be non-negative. We get this minimum value in the usual way by finding the value of t which makes the derivative of the function vanish. Thus

$$2(v\cdot v)t + 2(v\cdot w) = 0$$

and so

$$t = -v\cdot w/v\cdot v, \tag{2.4}$$

provided $v\cdot v \neq 0$. The exception can occur only if $v = 0$, in which case the theorem is trivially true. Hence we can substitute from (2.4) into (2.3) to get

$$0 \leqq (v\cdot v)\,\frac{(v\cdot w)^2}{(v\cdot v)^2} - 2(v\cdot w)\,\frac{v\cdot w}{v\cdot v} + w\cdot w$$

$$= [-\,(v\cdot w)^2 + (v\cdot v)(w\cdot w)]/v\cdot v\ .$$

Since $v\cdot v > 0$ this gives us the desired result.

Corollary 2.1. $|v\cdot w| \leqq \|v\|\,\|w\|$.

The triangle inequality is now easy to prove.

Theorem 2.4. $\|v + w\| \leqq \|v\| + \|w\|$.

Proof. If $\|v+w\| = 0$ the statement is obviously true. If $\|v+w\| \neq 0$, th

$$\begin{aligned}\|v + w\|^2 &= (v + w)\cdot(v + w)\\ &= (v + w)\cdot v + (v + w)\cdot w\\ &\leqq |(v + w)\cdot v| + |(v + w)\cdot w|\\ &\leqq \|v + w\|\,\|v\| + \|v + w\|\,\|w\|\\ &= \|v + w\|\,(\|v\| + \|w\|).\end{aligned}$$

The fourth step is obtained by applying Corollary 2.1. The desired result follows on dividing by $\|v + w\|$.

A vector of length 1 is called a unit vector. Any non-zero vector v has a unique positive scalar multiple of unit length, namely $\|v\|^{-1}v$, usually written $v/\|v\|$. The passage from v to $v/\|v\|$ is called <u>normalization</u>.

We can now give a definition of the angle between two vectors, and a corresponding definition of orthogonality.

<u>Definition 2.3</u>. If v and w are non-zero vectors we define the <u>angle</u> θ between them by means of

$$\cos\theta = \frac{v\cdot w}{\|v\| \ \|w\|}\ , \quad 0 \leq \theta \leq \pi$$

<u>Corollary 2.2</u>. The inner product of two unit vectors is the cosine of the angle between them.

By virture of Corollary 2.1, $|\cos\theta| \leqq 1$, and so θ can be chosen to be a real number in the range $0 \leqq \theta \leqq \pi$.

<u>Definition 2.4</u>. Two vectors v and w are said to be <u>orthogonal</u> if $v\cdot w = 0$.

This definition permits v or w to be zero, and says that $\vec{0}$ is orthogonal to any vector (including itself). The use of the term in this trivial case is convenient in avoiding special cases in such statements as the one in Theorem 2.5 below.

<u>Example 2.1a</u>. In V_n the condition for orthogonality of $v = (a_1,\ldots,a_n)$ and $w = (b_1,\ldots,b_n)$ is $\sum_{i=1}^{n} a_i b_i = 0$, and the angle between v and w is given by

$$\cos\theta = \frac{\sum a_i b_i}{\sqrt{\sum a_i^2 \sum b_i^2}}$$

(cf. Problem 1.1). The application of the Schwarz Inequality to this case gives the result known as the Cauchy Inequality:

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leqq \left(\sum_{i=1}^{n} a_i{}^2\right)\left(\sum_{i=1}^{n} b_i{}^2\right) .$$

Example 2.2a. In C^0, f and g are orthogonal if $\int_a^b f(x)g(x)dx = 0$; for example on the interval $0 \leqq x \leqq 1$ the two functions x and $x^2 - 1/2$ are orthogonal since

$$\int_0^1 x\left(x^2 - \frac{1}{2}\right)dx = \int_0^1 \left(x^3 - \frac{1}{2}x\right)dx = 0.$$

The length of the function $x^2 - 1/2$, considered as an element of this vector space, is

$$\left[\int_0^1 \left(x^2 - \frac{1}{2}\right)^2 dx\right]^{1/2} = \left[\int_0^1 \left(x^4 - x^2 + \frac{1}{4}\right)dx\right]^{1/2} = \sqrt{7/60} .$$

We close this section with one more useful theorem.

Theorem 2.5. If a vector is orthogonal to each of a given finite set of vectors then it is orthogonal to any linear combination of this set.

This follows immediately from Theorem 2.1(b).

Problems

2.1 Given the following vectors in V_3: $u = (1,2,3)$, $v = (3,2,1)$,

$-2 \quad 1/3 \quad 2/3$

$w = (-1,1,1)$.

$0 \quad \frac{5}{3} \quad \frac{4}{3} \qquad \frac{7}{2} \quad 4 \quad \frac{5}{3}$

(a) Compute each of the following:

$u \cdot v$, $\|u\|$, angle between u and w.

(b) Is any pair of the vectors orthogonal?

(c) Compute $(u \cdot v)w$, $u(v \cdot w)$. Are they equal?

(d) Find a linear combination of v and w which is orthogonal to u. What does this mean geometrically?

2.2 Show that as elements of C^0 on $-\pi \leqq x \leqq \pi$ the functions $\sin x$ and $\cos x$ are orthogonal.

2.3 As elements of C^0 on $0 \leqq x \leqq 1$, what is the angle in degrees between x and x^2? Answer. $\theta = 14.07^\circ$

2.4 Determine a and b so that in C^0 on $0 \leqq x \leqq 1$ the function $x^2 + ax + b$ is orthogonal to each of the functions 1 and x.

2.5 Prove Theorem 2.1, 2.2, and 2.5.

2.6 Prove that $u \cdot (av) = a(u \cdot v)$, $(u+v) \cdot w = u \cdot w + v \cdot w$, and $(u+v) \cdot (w+z) = u \cdot w + v \cdot w + u \cdot z + v \cdot z$.

2.7 Interpret the identity $(u-v) \cdot (u-v) = \|u\|^2 + \|v\|^2 - 2u \cdot v$ geometrically as the Law of Cosines.

2.8 If w is a fixed vector prove that the set of all vectors in V orthogonal to w is a vector subspace of V. Interpret this geometrically if V is V_3.

2.9 Show that multiplying vectors by non-zero scalars does not change their orthogonality properties, and that multiplying them by positive scalars does not change the angle between them. Hence angle and orthogonality are preserved when vectors are normalized.

2.10 (a) Prove that $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

(b) Prove that the sum of the squares of the diagonals of a parallelogram equals the sum of the squares of the sides, by interpreting (a) geometrically.

(c) Prove that $u \cdot v = (\|u+v\|^2 - \|u-v\|^2)/4 = (\|u+v\|^2 - \|u\|^2 - \|v\|^2)/2$.

2.11 Prove the Pythagorean Theorem for vector spaces, i.e., if u and v are orthogonal then $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

2.12 Prove that if $|u| = |v|$ then $u + v$ and $u - v$ are orthogonal. Interpret this geometrically.

2.13 Derive Cauchy's Inequality for functions: if $f(x)$ and $g(x)$ are continuous functions on the interval $a \leqq x \leqq b$ then

$$\left[\int_a^b f(x)g(x)dx\right]^2 \leqq \int_a^b f(x)^2 dx \int_a^b g(x)^2 dx.$$

2.14 Let $w(x)$ be a function which is continuous and positive on the interval $a \leqq x \leqq b$. Show that, like (2.2),

$$f \cdot g = \int_a^b w(x)f(x)g(x)dx$$

defines an inner product in C^0 on $a \leqq x \leqq b$. [$w(x)$ is called the <u>weight function</u> for this inner product.]

3. <u>Orthogonal Bases</u>.

At the end of Section 1 we saw that the most convenient basis to use in the 3-dimensional space of physical vectors was one whose members were pairwise orthogonal ($i \cdot j = i \cdot k = j \cdot k = 0$) and of unit lengths ($i \cdot i = j \cdot j = k \cdot k = 1$). Similarly, in dealing with an inner product in an abstract vector space V we shall find it useful to pay special attention to bases having similar properties.

<u>Definition 3.1</u>. A basis $\{v_1,\ldots,v_m\}$ of a vector space V is said to be <u>orthogonal</u> if $v_i \cdot v_j = 0$ for $i \neq j$. If in addition the v_i's are unit vectors, so that $v_i \cdot v_i = 1$, $i = 1,\ldots,m$, the basis is said to be <u>orthonormal</u>.

Since $\vec{0}$ is never an element of a basis, any orthogonal basis can be made into an orthonormal one by normalizing each of its members. Hence most of our attention will be concentrated on the orthogonality condition.

By definition a basis of a vector space is an independent set that spans the space. We now show that for an orthogonal basis the independence is automatically taken care of.

<u>Theorem 3.1</u>. Any set of non-zero pairwise orthogonal vectors is independent.

<u>Proof</u>. Let $v_1,\ldots,v_m$ be non-zero vectors such that $v_i . v_j = 0$ for $i \neq j$, and suppose that for some set of constants $a_1,\ldots,a_m$,

$$(3.1) \qquad a_1v_1 + \ldots + a_mv_m = 0.$$

Then

$$\begin{aligned} 0 = v_1 \cdot 0 &= v_1 \cdot (a_1v_1 + a_2v_2 + \ldots + a_mv_m) \\ &= a_1(v_1 \cdot v_1) + a_2(v_1 \cdot v_2) + \ldots + a_m(v_1 \cdot v_m) \\ &= a_1(v_1, v_1), \end{aligned}$$

since $v_1 \cdot v_i = 0$ if $i \neq 1$. Since $v_1 \neq 0$ we have $v_1 \cdot v_1 \neq 0$, and so $a_1 = 0$. In exactly the same fashion we can show that $a_2 = 0, \dots, a_m = 0$. Hence the only set $a_1, \dots, a_m$ for which (3.1) holds is the trivial set $a_1 = \dots = a_m = 0$, and so the set $\{v_1, \dots, v_m\}$ is independent.

Because of this theorem we can be sure that any set of pairwise orthogonal, non-zero vectors is a basis of the set it spans.

Example 3.1. The three vectors in V_4, $(1,1,-1,-1)$, $(1,-1,1,-1)$, $(1,-1,-1,1)$ are easily seen to form a pairwise orthogonal set. Hence they are an orthogonal basis for the 3-dimensional subspace W that they span. The length of each is 2, and so the corresponding orthonormal basis of W is

$$\left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right).$$

V_n itself has an obvious orthonormal basis, namely $\{\vec{e}_1, \cdots, \vec{e}_n\}$, where $\vec{e}_i$ is the n-tuple whose i-th element is 1 and all others are 0. But the above example indicates that we might sometimes be interested in subspaces of V_n, or perhaps finite-dimensional subspaces of C^0, for which an orthogonal basis is not at all obvious. We thus are led to the problem: Given an independent set of vectors $\{v_1, \dots, v_m\}$ in a vector space V, spanning a subspace W, how can we determine from them an orthogonal basis for W? The most common method of doing this is known as the Orthogonalization Process of Schmidt.

Theorem 3.2. Let $\{v_1,\ldots,v_m\}$ be an independent set of vectors spanning a space W. Define successively,

$$w_1 = v_1,$$

$$w_2 = v_2 - \frac{v_2 \cdot w_1}{w_1 \cdot w_1} w_1,$$

$$w_3 = v_3 - \frac{v_3 \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{v_3 \cdot w_2}{w_2 \cdot w_2} w_2,$$

. . . .

$$w_m = v_m - \frac{v_m \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{v_m \cdot w_2}{w_2 \cdot w_2} w_2 - \cdots - \frac{v_m \cdot w_{m-1}}{w_{m-1} \cdot w_{m-1}} w_{m-1} .$$

Then $\{w_1,\ldots,w_m\}$ is an orthogonal basis for W.

Proof of this theorem is given at the end of this section.

Example 3.2. In V_4 find an orthonormal basis for the subspace W spanned by

$$\{(1,2,3,4),\ (4,3,2,1),\ (0,1,0,1)\}.$$

For hand computation we might as well start with the simplest vector, so we take

$$v_1 = (0,1,0,1),\ v_2 = (1,2,3,4),\ v_3 = (4,3,2,1).$$

Then

$$w_1 = (0,1,0,1);$$

$$w_2 = (1,2,3,4) - \frac{6}{2}(0,1,0,1) = (1,-1,3,1);$$

$$w_3 = (4,3,2,1) - \frac{4}{2}(0,1,0,1) - \frac{8}{12}(1,-1,3,1) = (\frac{10}{3}, \frac{5}{3}, 0, -\frac{5}{3})$$

Normalizing these gives the orthonormal basis for W:

$$\{w_1', w_2', w_3'\} = \left\{\left(0,\ \frac{1}{\sqrt{2}},\ 0,\ \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{12}},\ -\frac{1}{\sqrt{12}},\ \frac{3}{\sqrt{12}},\ \frac{1}{\sqrt{12}}\right),\right.$$

$$\left.\left(\frac{2}{\sqrt{6}},\ \frac{1}{\sqrt{6}},\ 0,\ -\frac{1}{\sqrt{6}}\right)\right\}.$$

<u>Example 3.3.</u> In C^0 on $0 \leqq x \leqq 1$ find an orthogonal basis for the subspace spanned by $\{1, x, x^2, \ldots, x^n\}$.

$$w_1 = 1;$$

$$w_2 = x - \frac{\int_0^1 (x)(1)dx}{\int_0^1 1^2 dx}(1) = x - \frac{1}{2};$$

$$w_3 = x^2 - \frac{\int_0^1 (x^2)(1)dx}{\int_0^1 1^2 dx}(1) - \frac{\int_0^1 (x^2)(x - \frac{1}{2})dx}{\int_0^1 (x - \frac{1}{2})^2 dx}(x - \tfrac{1}{2})$$

$$= x^2 - \frac{1}{3} - (x - \tfrac{1}{2}) = x^2 - x + \frac{1}{6};$$

$$w_4 = x^3 - \frac{\int_0^1 x^3 dx}{\int_0^1 1^2 dx}(1) - \frac{\int_0^1 x^3(x - \frac{1}{2})dx}{\int_0^1 (x - \frac{1}{2})^2 dx}(x - \tfrac{1}{2})$$

$$- \frac{\int_0^1 x^3(x^2 - x + \frac{1}{6})dx}{\int_0^1 (x^2 - x + \frac{1}{6})^2 dx}(x^2 - x + \tfrac{1}{6})$$

$$= x^3 - \frac{1}{4} - \frac{9}{10}(x - \tfrac{1}{2}) - \frac{3}{2}(x^2 - x + \tfrac{1}{6})$$

$$= x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20};$$

and so on. When we refer to "the orthogonal polynomials" on the interval $0 \leqq x \leqq 1$ we mean those obtained by Schmidt orthogonalization of $\{1, x, x^2, \ldots\}$.*

By definition, if $\{v_1, \ldots, v_m\}$ is a basis for W, and v is any vector in W we can express v as a linear combination of basis vectors.

$$v = \sum_{i=1}^{m} a_i v_i \quad . \tag{3.2}$$

Let $w = \sum_{j=1}^{m} b_j v_j$ be another vector in W. Then(compare with the discussion at the end of Section 1)

$$v \cdot w = \left(\sum_{i=1}^{m} a_i v_i \right) \cdot \left(\sum_{j=1}^{m} b_j \; v_j \right) ,$$

or

$$v \cdot w = \sum_{i=1}^{m} \; \sum_{j=1}^{m} a_i b_j (v_i \cdot v_j) \quad .$$

If the basis is orthonormal we have

$$v_i \cdot v_j = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases} , \tag{3.3}$$

and (3.3) reduces to

$$v \cdot w = \sum_{i=1}^{m} a_i b_i \quad .$$

---

*If a weight function (cf. Problem 2.14) is used, the Schmidt process produces a different set of orthogonal polynomials. We will use only the ones defined here.

Thus a general orthonormal basis behaves with regard to inner product exactly like the special basis $\{\vec{e}_1,\ldots,\vec{e}_n\}$ of V_n.

Now we must consider the problem of expressing a vector in terms of a given basis. In Example 5.5 of Chapter 2 a method was introduced which, for vectors in V_n, reduces this problem to the solution of a system of n linear equations. This technique is satisfactory if n is not too large, but for many purposes a more direct method is desirable. The following theorem gives such a method for the case in which the basis is orthogonal. It is not restricted to V_n but works in any vector space.

Theorem 3.3. If v is a vector in the space spanned by the orthogonal basis $\{v_1,\ldots,v_m\}$, then

$$v = \sum_{i=1}^{m} a_i v_i,$$

with

$$a_i = \frac{v \cdot v_i}{v_i \cdot v_i}, \quad i = 1,\ldots,m \ . \tag{3.5}$$

Proof. We have

$$\begin{aligned} v\cdot v_1 &= (a_1v_1 + a_2v_2 + \ldots + a_mv_m)\cdot v_1 \\ &= a(v_1\cdot v_1) + a_2(v_2\cdot v_1) + \ldots + a_m(v_m\cdot v_1) \\ &= a_1(v_1\cdot v_1), \end{aligned}$$

since in the orthogonal basis we have $v_j\cdot v_1 = 0$ if $j \neq 1$.

This gives (3.5) for i = 1, and the result for any other value of i is obtained similarly.

Corollary 3.1. If $\{v_1, \ldots, v_m\}$ is an orthonormal basis for a subspace W, then any vector v in W is given by $V_i \cdot V_i = 1$
(orthonormality!)

$$v = \sum_{i=1}^{m} (v \cdot v_i) v_i, \quad \text{and} \quad \|v\|^2 = \sum_{i=1}^{m} (v \cdot v_i)^2.$$

Example 3.2, continued. Let us express the vector v = (1,1,1,1), which lies in the subspace W -- since $(1,1,1,1) = \frac{1}{5}((1,2,3,4) + (4,3,2,1))$ -- in terms of the orthonormal basis $\{w_1', w_2', w_3'\}$. We have

$$v \cdot w_1' = \frac{2}{\sqrt{2}}, \; v \cdot w_2' = \frac{4}{\sqrt{12}}, \; v \cdot w_3' = \frac{2}{\sqrt{6}} \;.$$

Hence

$$v = \frac{2}{\sqrt{2}} w_1' + \frac{4}{\sqrt{12}} w_2' + \frac{2}{\sqrt{6}} w_3'.$$

This example prompts us to ask a question. Suppose v does not lie in the subspace spanned by the given orthogonal basis, what do we get if we apply the process of Theorem 3.3? The answer is given by the following theorem.

Theorem 3.4. If $\{v_1, \ldots, v_m\}$ is an orthogonal basis for a subspace W of V, if v is any vector in V, and if w is defined by

$$w = \sum_{i=1}^{m} a_i v_i, \tag{3.6}$$

where

$$a_i = \frac{v \cdot v_i}{v_i \cdot v_i}, \qquad i = 1, \ldots, m \quad , \tag{3.7}$$

then $v = w + u$, where w is in W and u is a vector in V orthogonal to every vector in W.

Proof. Obviously w lies in W, since by (3.6) it is a linear combination of the basis of W. If we apply the previous theorem to w we find that

$$a_i = \frac{w \cdot v_i}{v_i \cdot v_i} . \tag{3.8}$$

From (3.7) and (3.8) we conclude that

$$w \cdot v_i = v \cdot v_i, \quad i = 1, \ldots, m \quad . \tag{3.9}$$

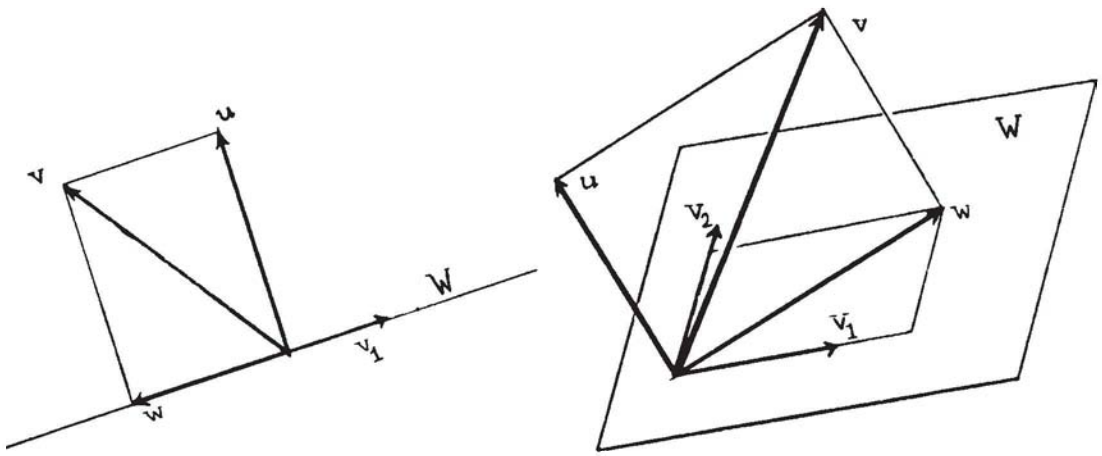
Now let $u = v - w$. Then

$$u \cdot v_i = v \cdot v_i - w \cdot v_i = 0, \quad i = 1, \ldots, m ,$$

from (3.9). Hence u is orthogonal to every v_i, and so from Theorem 2.5 to every vector in W.

The vector w defined by (3.6) and (3.7) is called the orthogonal projection of v into W, and u is the component of v orthogonal to W. These terms have simple geometric interpretations in V_3 (Figure 3.1).

The orthogonal projection of a vector into a subspace has a property which is very useful in certain problems of approximation. This property is expressed in the following theorem:



W is 1-dimensional W is 2-dimensional

Figure 3.1

Theorem 3.5. Given a vector v and a subspace W. Of all vectors y in W the one that is closest to v, in the sense that $\|v-y\|$ is a minimum, is the orthogonal projection of v into W.

Proof. Let w be the orthogonal projection of v into W; then $v = w+u$, where u is orthogonal to every vector in W (Theorem 3.4). Then

$$v - y = w + u - y = u + z$$

where $z = w - y$ is a vector in W since both w and y are in W. Hence

$$\begin{aligned} \| v - y \|^2 &= \| u + z \|^2 \\ &= (u + z)\cdot(u + z) \\ &= u\cdot u + 2u\cdot z + z\cdot z \\ &= \| u \|^2 + \| z \|^2 , \end{aligned}$$

the term $2u \cdot z$ being zero since z is in W and u is orthogonal to every vector in W. Replacing z by its equal w - y gives

$$\| v - y \| = [\| u \|^2 + \| w - y \|^2]^{1/2}$$

and quite obviously the choice of y that makes $\| v - y \|$ a minimum is y = w, since for any other y the term $\| w - y \|^2$ is positive. This is what we wished to prove.

Example 3.4. In constructing a computer compiling program for a programming language such as CORC we must provide the machine with means for quickly computing the values of such elementary functions as, for example, e^x. This is frequently done by approximating the function over some interval by a suitable polynomial and using special properties of the function to extend the approximation to values of x outside of this interval. If, for instance, we approximate e^x over the interval $0 \leqq x \leqq 1$ by a third degree polynomial, thus

$$e^x \cong a + bx + cx^2 + dx^3 , \tag{3.10}$$

we can compute $e^{3.1416}$ by

$$e^{3.1416} = e^3 e^{.1416} ,$$

getting e^3 by direct multiplication and $e^{.1416}$ from (3.10). Our problem then reduces to the selection of the constants a,b,c,d in (3.10).

We have already considered one approach in Chapter 3. We can expand e^x in a Maclaurin Series and take the first four log(1+x) = $\frac{1}{1+x}$ log(1+x) = $\frac{1}{1+x}$ $\frac{1}{(1+x)^2}$

terms, thus

$$e^x \cong 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 . \tag{3.11}$$

This gives a very good approximation for values of x near zero, but a poor one for x near 1. A better fit is obtainable by expanding in a Taylor Series about the mid-point of the interval, thus

$$\begin{aligned} e^x = e^{1/2}e^{x-1/2} &\cong e^{1/2}(1 + (x - \tfrac{1}{2}) + \tfrac{1}{2}(x - \tfrac{1}{2})^2 + \tfrac{1}{6}(x - \tfrac{1}{2})^3) \\ &= e^{1/2}(\tfrac{29}{48} + \tfrac{5}{8}\,x + \tfrac{1}{4}\,x^2 + \tfrac{1}{6}\,x^3) \\ &= 0.99610 + 1.03045x + 0.41218x^2 + 0.27479x^3. \end{aligned} \tag{3.12}$$

Neither of these, however, could be called a "closest" approximation. To get something of this sort we apply Theorem 3.5 to the vector space C^0 over the interval $0 \leqq x \leqq 1$. Note that e^x is a vector in C^0 and the vectors of the form $a + bx + cx^2 + dx^3$ are linear combinations of the independent vectors $1, x, x^2, x^3$ and hence constitute a 4-dimensional subspace W with these four vectors as a basis. Thus e^x plays the role of v in Theorem 3.5 and $a + bx + cx^2 + dx^3$ the role of y. The difference v-y is just the error E in the approximation (3.10) and the theorem tells us how to minimize

$$\| E \| = [\int_0^1 E^2 dx]^{1/2} .$$

This expression is called the "root mean square" of the error.

Since it depends on values of the function e^x over the whole interval from 0 to 1 we can expect to get better results from this approach than from the power series, which depend on the values of the function and its derivatives at only one point of the interval.

We wish, then, to construct the projection of e^x into the subspace W with basis $1, x, x^2, x^3$. The machinery for this is given in Theorem 3.4. The first step is to orthogonalize the basis by using the Schmidt Process. This is exactly what was done in Example 3.3. The orthogonal basis is

$$w_1 = 1, \qquad w_1 \cdot w_1 = 1,$$

$$w_2 = x - \frac{1}{2}, \qquad w_2 \cdot w_2 = 1/12,$$

$$w_3 = x^2 - x + \frac{1}{6}, \qquad w_3 \cdot w_3 = 1/180,$$

$$w_4 = x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}, \qquad w_4 \cdot w_4 = 1/2800.$$

The first three values of $w_i \cdot w_i$ are obtained in the course of the orthogonalization; the last must be computed separately. We can now apply Theorem 3.4, with $v = e^x$, as follows:

$$a_1 = \frac{v \cdot w_1}{w_1 \cdot w_1} = \int_0^1 e^x dx = e-1 = 1.71828,$$

$$a_2 = \frac{v \cdot w_2}{w_2 \cdot w_2} = 12\int_0^1 e^x(x - 1/2)dx = 12\ (\tfrac{3}{2} - \tfrac{1}{2}e) = 1.69031,$$

$$a_3 = \frac{v \cdot w_3}{w_3 \cdot w_3} = 180\int_0^1 e^x(x^2-x+1/6)dx = 180\ (\tfrac{7}{6}e - \tfrac{19}{6}) = 0.83918,$$

$$a_4 = \frac{v \cdot w_4}{w_4 \cdot w_4} = 2800\int_0^1 e^x(x^3-(3/2)x^2 + (3/5)x-1/20)dx$$

$$= 2800(\frac{193}{20} - \frac{71}{20} e) = 0.27863.$$

The orthogonal projection w is then $w = 1.71828 + 1.69031(x - 1/2) + 0.83918(x^2-x+1/6) + 0.27863(x^3-(3/2)x^2 + (3/5)x - 1/20)$, and so (3.10) becomes

(3.13) $e^x \cong 0.99906 + 1.01831x + 0.42124\, x^2 + 0.27863x^3.$

To see how (3.11), (3.12) and (3.13) compare, we have computed below the values at x = 0, 0.5 and 1 obtained from these formulas, along with the values of e^x correct to five places

x	0	0.5	1
e^x	1.00000	1.64872	2.71828
(3.11)	1.00000	1.64583	2.66667
(3.12)	0.99610	1.64872	2.71351
(3.13)	0.99906	1.64835	2.71724

(3.13) is correct to within about .001 throughout the interval, (3.12) is off by nearly .005 at x = 1, and (3.11) is off by .05 at x = 1. Thus for a four-term approximation to e^x the root-mean-square method is superior to either power series method, and this is typical for most polynomial approximations over an interval.

Proof of Theorem 3.2. For convenience in the following discussion we are going to use a_{ij} as an abbreviation for $\frac{v_i \cdot w_j}{w_j \cdot w_j}$, so that the equations of Theorem 3.2 take the more easily handled form

$a_{ij} = \frac{v_i w_j}{w_j w_j}$

$$w_1 = v_1,$$

$$w_2 = v_2 - a_{21}w_1,$$

$$w_3 = v_3 - a_{31}w_1 - a_{32}w_2,$$

. . . .

$$w_m = v_m - a_{m1}w_1 - a_{m2}w_2 - \dots - a_{m,m-1}w_{m-1}.$$

We must prove two things:

(1) No w_i is zero. This enables us to divide by $w_i \cdot w_i$ in forming $a_{i+1,i}$ needed in the next step.

(2) Each w_i is orthogonal to all the preceding w's.

The proof of (1) is by reductio ad absurdum. Assume that some of the w's are zero, and take the one with smallest subscript; suppose, for definiteness, that this is w_3. Then $w_1 \neq 0$, $w_2 \neq 0$, $w_3 = 0$. Then we have

$$w_1 = v_1 \neq 0$$

$$w_2 = v_2 - a_{21}w_1 \neq 0$$

$$0 = v_3 - a_{31}w_1 - a_{32}w_2.$$

The value of w_1 from the first equation can be substituted into the second. Then the values of w_1 and w_2 from these two can be substituted into the third, giving a non-trivial linear relation between v_1, v_2 and v_3 (non-trivial since the coefficient of v_3 is 1). This is impossible since the v's are given as independent. Hence we cannot have $w_3 = 0$; and similarly we cannot have $w_i = 0$ for any i.

(2). Here we proceed step-by-step. (Or a more elegant proof can be given by mathematical induction.) Note that by definition of a_{ij} we have

$$a_{ij}w_j \cdot w_j = v_i \cdot w_j \ .$$

We will refer to the above equation as (*ij); e.g., (*21) refers to the equation

$$a_{21}w_1 \cdot w_1 = v_2 \cdot w_1 \ .$$

Step 1.

$$\begin{aligned} w_2 \cdot w_1 &= (v_2 - a_{21}w_1) \cdot w_1 \\ &= v_2 \cdot w_1 - a_{21}w_1 \cdot w_1 \\ &= v_2 \cdot w_1 - v_2 \cdot w_1 = 0, \end{aligned}$$

by (*21).

Step 2a.

$$\begin{aligned} w_3 \cdot w_1 &= (v_3 - a_{31}w_1 - a_{32}w_2) \cdot w_1 \\ &= v_3 \cdot w_1 - a_{31}w_1 \cdot w_1 - a_{32}w_2 \cdot w_1 \\ &= v_3 \cdot w_1 - v_3 \cdot w_1 - a_{32}0 = 0, \end{aligned}$$

by (*31) and Step 1.

Step 2b.

$$\begin{aligned} w_3 \cdot w_2 &= (v_3 - a_{31}w_1 - a_{32}w_2) \cdot w_2 \\ &= v_3 \cdot w_2 - a_{31}w_1 \cdot w_2 - a_{32}w_2 \cdot w_2 \\ &= v_3 \cdot w_2 - a_{31}0 - v_3 \cdot w_2 = 0, \end{aligned}$$

by Step 1 and (*32).

It is easy to see how the proof continues.

Problems

3.1. Show that $\{(1,2,2), (2,1,-2), (2,-2,1)\}$ is an orthogonal basis for V_3. Normalize it.

3.2. Express (1,2,3) in terms of the orthogonal (or orthonormal) basis of Problem 3.1.

3.3. Use the Schmidt Process to get an orthogonal basis for the subspace W of V_4 spanned by $\{(1,1,1,1), (1,2,3,4), (-4,3,-2,1)\}$. Find a vector orthogonal to W.

3.4. Find the first four orthogonal polynomials on the interval $-1 \leqq x \leqq 1$. Answer. $1, x, x^2 - 1/3,\ x^3 - 3x/5$.

3.5. Let W be the subspace of Example 3.1. For each of the following vectors find its orthogonal projection into W and its component orthogonal to W;

(a) (1,2,3,4),

(b) (1,-2,-3,4),

(c) (1,1,1,1).

Answers. (a) (1/2)(-3,-1,1,3),(5/2)(1,1,1,1);(c) (0,0,0,0),(1,1,1,1).

3.6. In Theorem 3.4 the decomposition of the vector v into parts w and u respectively in W and orthogonal to W was made in terms of a given orthogonal basis of W. Show that such a decomposition does not depend on the particular basis by proving the following:

Let W be a linear subspace of V and let v be any vector in V. Let $v = w_1 + u_1$ and also $v = w_2 + u_2$, where w_1 and w_2 are in W and u_1 and u_2 are each orthogonal to W. Then necessarily $w_1 = w_2$ and u_1 and u_2; that is, there is at most one such decomposition.

[Hint: Let $z = w_1 - w_2 = u_2 - u_1$ and show that $z \cdot z = 0$.]

3.7. With the notation of Theorem 3.4 prove that

$$\| u \|^2 + \sum_{i=1}^{m} \frac{(v \cdot v_i)^2}{\| v_i \|^2} = \| v \|^2,$$

and consequently that

log (1+x)

$\frac{1}{1+x}$ $a + bx - cx^2 + dx^3$

$$\sum_{i=1}^{m} \frac{(v \cdot v_i)^2}{\| v_i \|^2} \leqq \| v \|^2 .$$

3.8. Carry out Step 3 of the second part of the proof of Theorem 3.2.

3.9. Use the method of Example 3.4 to get the root-mean-square approximation to log (1 + x) on the interval $0 \leqq x \leqq 1$ by a third degree polynomial. The following steps are suggested to simplify the work.

(a) Using integration by parts followed by a substitution derive the formula

$$c_n = \int_0^1 x^n \log(1+x)\,dx = \frac{1}{n+1}\left(\log 2 - \int_1^2 \frac{(t-1)^{n+1}}{t}\,dt\right).$$

(b) Compute c_0, c_1, c_2, c_3. (Note. log 2 = 0.69314718.)

(c) From c_0, c_1, c_2, c_3 compute the a_1, a_2, a_3, a_4 of Theorem 3.4.

(d) Get the root-mean-square approximation to log (1+x).

(e) Compare the approximation in (d) with the accurate value of log (1+x) at x = 0, 0.5, 1.

Answers:

(b) $c_0 = 2 \log 2 - 1 = .38629436$; $c_1 = \frac{1}{4} = .25$;

$c_2 = \frac{2}{3} \log 2 - \frac{5}{18} = .18432034$; $c_3 = \frac{7}{48} = .14583333$.

(c) $a_1 = .38629$; $a_2 = .68223$; $a_3 = -.23351$; $a_4 = .01067$.

(d) $.00572 + .92214x - .24952x^2 + .01067x^3$.

3.10. Show that the orthogonal polynomials on a given interval $a \leqq x \leqq b$ have the following properties:

(a) The (n+1)-th orthogonal polynomial $P_n(x)$ is a polynomial of degree n whose highest term is x^n.

(b) Any polynomial of degree at most n is expressible as a linear combination of orthogonal polynomials of degrees at most n.

(c) If Q(x) is any polynomial of degree less than n, then

$$\int_a^b P_n(x)Q(x)dx = 0 .$$

(d) In any integral of the form $\int_a^b P_n(x)Q(x)dx$ the terms of Q(x) of degrees less than n can be ignored. In particular

$$\int_a^b P_n(x)^2 dx = \int_a^b x^n P_n(x)dx .$$

(e) Let

$$\int_a^b P_n(x)^2 dx = \rho_n, \quad n = 0,1,2,\ldots .$$

Then if Q(x) is any polynomial of degree m we have

$$Q(x) = \sum_{n=0}^{m} \frac{1}{\rho_n} P_n(x) \int_a^b P_n(x)Q(x)dx \ .$$

4. Matrices. An Application.

In Chapter 2 we found that in working with vectors and matrices in the vector space V_n it is convenient to consider a vector as an n x 1 matrix,

$$v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} ,$$

or a column vector. A row vector, a 1 x n matrix, can then be regarded as the transpose v^t of the corresponding column vector. Using these conventions it is easy to verify that the inner product $u \cdot v$ of two vectors in V_n is just the matrix product $u^t v$, or, equally well, $v^t u$. Thus the inner product is brought within the realm of matrix algebra.

To give an application of the use of matrix algebra, and also to provide motivation for a later topic, we consider a problem in structural mechanics.

Many structures such as bridges, steel buildings, etc. are made by fastening together units known as "members." In analyzing such structures it is often appropriate to assume that these members are "pin-jointed." This means that members are joined only at the ends and there are no torques acting at the joints. We also frequently assume that the weights of the members are negligible. Thus for mathematical analysis the members can be considered to be weightless line segments, with forces acting only at the ends and along the segments. In some cases (the so-called "statically determinate" cases) this is all that need be assumed. But to give an interesting application of matrix algebra we wish to discuss the statically indeterminate case and consider the elastic reaction of a member. This is merely Hooke's Law; that is, the tension in a member is proportional to the amount of stretching. Since this is a very crude model of the true physical picture, we are justified in making approximations to simplify our mathematics.

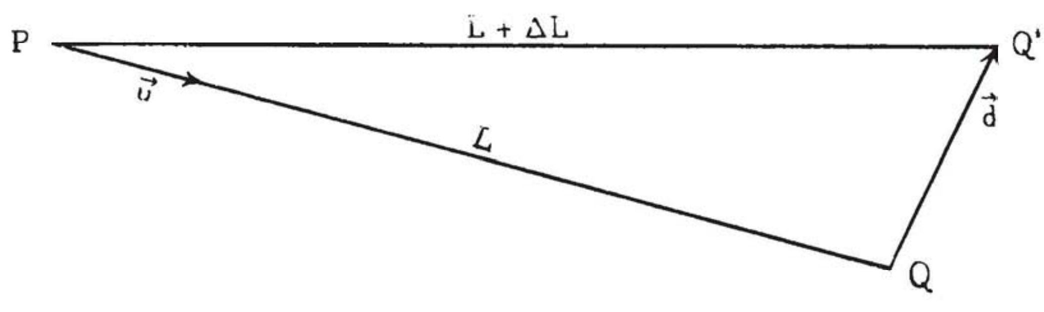


Figure 4.1

Figure 4.1 shows a member PQ of length L whose one end P is fixed but whose other end Q is displaced a short distance from Q to Q'. This displacement changes the length of the member from L to L + ΔL and so induces a tension (or compression if $\Delta L < 0$) in it. We want to find an expression for this tension in terms of the displacement.

Let $\vec{d}$ be the vector displacement and let $\vec{u}$ be the unit vector in the direction from P to Q. Then the vector from P to Q' is $L\vec{u} + \vec{d}$, and

$$(L + \Delta L)^2 = (L\vec{u} + \vec{d})\cdot(L\vec{u} + \vec{d}) \ ,$$

or

$$L^2 + 2L(\Delta L) + (\Delta L)^2 = L^2\vec{u}\cdot\vec{u} + 2L\vec{u}\cdot\vec{d} + \vec{d}\cdot\vec{d} \ . \tag{4.1}$$

Since $\vec{u}$ is a unit vector $\vec{u}\cdot\vec{u} = 1$. Also, since we assumed that the displacement was very small, we can make the approximation of neglecting the terms $\vec{d}\cdot\vec{d}$ and $(\Delta L)^2$. This type of approximation is very common in applied mathematics and is often referred to as "linearizing the equations" or "neglecting the higher order terms." In most cases it leads to meaningful results, but occasionally it does not. In these cases one should be able to trace back and find the step that causes the trouble.

With this approximation (4.1) reduces to the very simple form

$$\Delta L = \vec{u}\cdot\vec{d} \ . \tag{4.2}$$

Applying Hooke's Law gives us

(4.3) $$T = k\vec{u}\cdot\vec{d} \ ,$$

where T is the tension in the member and k an appropriate elastic constant for the member.

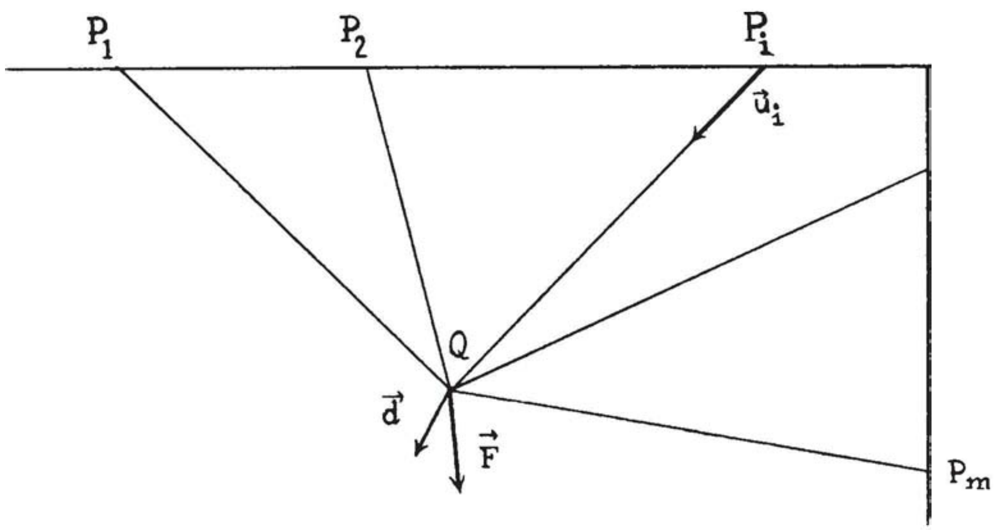


Figure 4.2

Figure 4.2 represents a simple structure with one moveable point Q connected by m pin-jointed members to fixed points $P_1, P_2, \ldots, P_m$. We apply a force $\vec{F}$ to the structure at Q and ask what is the displacement $\vec{d}$ of Q.

To solve this problem we reverse the point of view, supposing the displacement $\vec{d}$ known and considering the tension in each member caused by the displacement. If $\vec{u}_i$ is the unit vector from P_i to Q, and k_i the elastic constant for the corresponding member, this tension is, from (4.3),

$$T_i = k_i \vec{u}_i \cdot \vec{d} .$$

The component of $\vec{F}$ needed to balance this must have this magnitude and must act in the direction of $\vec{u}_i$; in other words, we must have

$$\vec{F} = \sum_{i=1}^{m} k_i (\vec{u}_i \cdot \vec{d}) \vec{u}_i . \tag{4.4}$$

(Note. In deriving this we have neglected the slight change in direction of the $\vec{u}_i$ due to the small displacement $\vec{d}$. A careful analysis shows that this leads to "higher order terms," which are not important in the approximation we are making.)

Equation (4.4) gives the basic relation between the applied force $\vec{F}$ and the displacement $\vec{d}$. The scalars k_i are presumably known from physical properties of the members and the vectors $\vec{u}_i$ from the geometry of the structure. Unfortunately one is usually interested in finding $\vec{d}$ on being given $\vec{F}$, and (4.4) does not look very promising from that point of view. To see how to handle this problem we transfer to matrix notation

If each $\vec{u}_i$ is a column vector,

$$\vec{u}_i = \begin{pmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \end{pmatrix}$$

then the set of all u_i, $i = 1,\ldots,m$, can be regarded as a $3 \times m$ matrix $U = (u_{ji})$. $\vec{d}$ and $\vec{F}$ become simply the column vectors,

$$d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} , \quad F = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} .$$

The scalars $k_1,\dots,k_m$ behave in an unusual fasnion; to handle them it will be convenient to introduce the diagonal matrix

$$K = \begin{pmatrix} k_1 & & & 0 \\ & k_2 & & \\ & & \ddots & \\ 0 & & & k_m \end{pmatrix} ,$$

an m x m matrix having these elements on its diagonal and zeros everywhere else.

We now have the machinery to proceed. In terms of vector components (4.4) can be written

$$\begin{aligned} f_h &= \sum_{i=1}^{m} k_i \left(\sum_{j=1}^{3} u_{ji} d_j \right) u_{hi} \\ &= \sum_{i=1}^{m} \sum_{j=1}^{3} u_{ji} k_i u_{hi} d_j \\ &= \sum_{j=1}^{3} \left(\sum_{i=1}^{m} u_{hi} k_i u_{ji} \right) d_j \\ &= \sum_{j=1}^{3} \left(\sum_{i=1}^{m} u_{hi} k_i u_{ij}{}^{t} \right) d_j \; . \end{aligned}$$

Now $u_{hi}k_i$ is easily seen to be the (hi)-th element of the matrix product UK, and so the inner sum is just the (hj)-th element of UKU^t. Therefore we may finally write (4.4) in the matrix form

(4.5) $$F = Md,$$

where

(4.6) $$M = UKU^t.$$

The use we make of (4.5) depends on the nature of the problem we are considering. If we want to find the force necessary to produce a given displacement , (4.5) gives us the answer directly. If we want the displacement produced by a given force, a more common problem, we must solve (4.5) for d. This can always be done if M is non-singular, using Gauss Elimination, Cramer's Rule, or any other convenient method. If we want to consider several loads F for the same structure it may be worth while to compute M^{-1} once for all and write (4.5) in the form

(4.7) $$d = M^{-1}F .$$

Once we know d we can, if we wish, find the tension in any of the members by using (4.3).

The situation when M is singular is worth investigating In this case the equation $Md = 0$ has a non-trivial solution. This means that there is a displacement d which, to a first approximation, can be effected without applying any force to the structure. A structure of this type is said to be <u>non-rigid</u>. A trivial example of such a "structure" is just a single bar pivoted at one end. Here $m = 1$, and M is of rank 1 and necessarily singular. A less trivial example is illustrated in Figure 4.3. This is physically rigid, but the rigidity depends on "second-order terms" in the displacement, and in our approximation these have been ignored. As far as our linear analysis is concerned a horizontal displacement of Q requires no force.

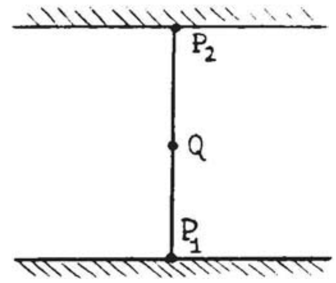
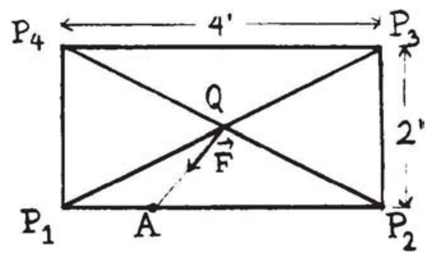
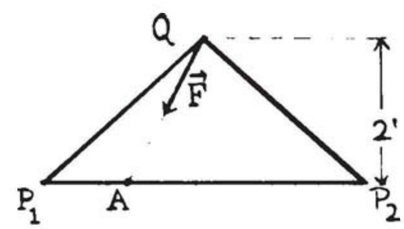


Figure 4.3

Example 4.1. Four 3-foot bars form the edges of a pyramid with base 2'x4' and altitude 2' (Figure 4.4). The corners of the base, P_1, P_2, P_3, P_4, are anchored, and at the vertex Q there is applied a force $\vec{F}$ directed towards a point A on the long side of the base one foot from the corner P_1. Each of the bars will stretch 0.1 inch under a tension of 1000 pounds. Find the approximate displacement of Q if the magnitude of $\vec{F}$ is one ton.



Top View



Front View

Figure 4.4

We first choose a rectangular coordinate system with origin at P_1, x - axis along P_1P_2, y - axis along P_1P_4 and z - axis upward. Then the coordinates of the various points are:

$P_1(0,0,0)$, $P_2(4,0,0)$, $P_3(4,2,0)$, $P_4(0,2,0)$, $Q(2,1,2)$, $A(1,0,0)$

From these we get the unit vectors

$$u_1 = \frac{1}{3}\begin{pmatrix}2\\1\\2\end{pmatrix},\ u_2 = \frac{1}{3}\begin{pmatrix}-2\\1\\2\end{pmatrix},\ u_3 = \frac{1}{3}\begin{pmatrix}-2\\-1\\2\end{pmatrix},\ u_4 = \frac{1}{3}\begin{pmatrix}2\\-1\\2\end{pmatrix},\ \text{and } F = \frac{2000}{\sqrt{6}}\begin{pmatrix}-1\\-1\\-2\end{pmatrix}$$

The matrix U is

$$U = \frac{1}{3}\begin{pmatrix}2 & -2 & -2 & 2\\1 & 1 & -1 & -1\\2 & 2 & 2 & 2\end{pmatrix} .$$

For each bar the constant k in (4.3) is given by

$$1000 = k\,\frac{0.1}{12}\ , \quad \text{or } k = 120{,}000 .$$

The diagonal matrix K is then just 120,000 I, where I is the 4 x 4 unit matrix.

Hence

$$M = UKU^t = 120{,}000\ UU^t$$

$$= \frac{120{,}000}{9}\begin{pmatrix}2 & -2 & -2 & 2\\1 & 1 & -1 & -1\\2 & 2 & 2 & 2\end{pmatrix}\begin{pmatrix}2 & 1 & 2\\-2 & 1 & 2\\-2 & -1 & 2\\2 & -1 & 2\end{pmatrix}$$

$$= \frac{40{,}000}{3}\begin{pmatrix}16 & 0 & 0\\0 & 4 & 0\\0 & 0 & 16\end{pmatrix} ,$$

and

$$M^{-1} = \frac{3}{640{,}000}\begin{pmatrix}1 & 0 & 0\\0 & 4 & 0\\0 & 0 & 1\end{pmatrix} .$$

Finally,

$$d = M^{-1}F = \frac{3}{320\sqrt{6}}\begin{pmatrix}-1\\-4\\-2\end{pmatrix} = .0175u,$$

where u is the unit vector $\frac{1}{\sqrt{21}}\begin{pmatrix}-1\\-4\\-2\end{pmatrix}$.

Thus the displacement is .0175 feet = .21 inch in the direction (-1, -4, -2).

Problems

4.1 A hook Q is attached to a wall and a ceiling by four pin-jointed members as shown. The members are uniform in the sense that the force required to stretch any of them varies inversely as the length and is 10,000 lbs per in. for a member one foot long. Find the displacement of Q if a load of one ton is hung on the hook.

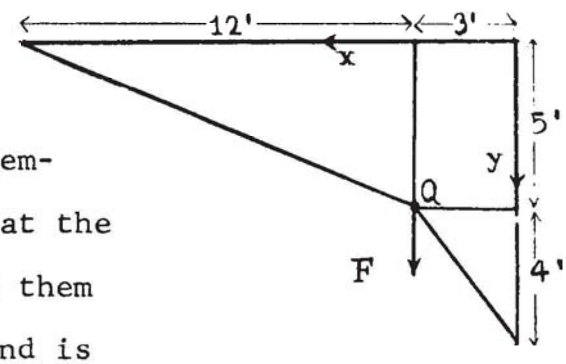


Figure 4.5

Answer. With axes as indicated,

$$d = \begin{pmatrix}.17\\.65\end{pmatrix} \text{ in.}$$

4.2 Figure 4.6 is a top view of a pin-jointed structure consisting of three members joining a point Q, directly above P_2, to points P_1, P_2 and P_3. These three points lie in a horizontal plane four feet below Q. The members are of the type described in Problem 4.1. At Q there is applied a horizontal force F of 1000 lbs. in the direction indicated. Find the displacement of Q and the stresses in the three members.

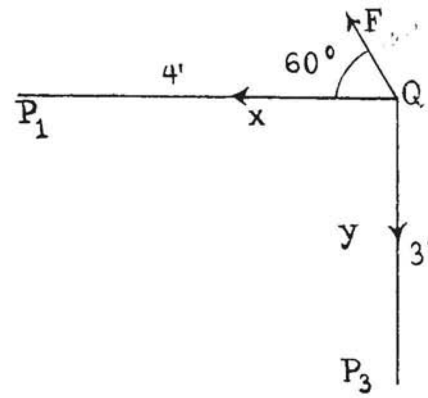


Figure 4.6

Answer. With the z-axis upward,

$$d = \begin{pmatrix} 0.30 \\ -1.55 \\ 0.26 \end{pmatrix} \text{ inches, } T_1 \cong -700 \text{ lb. (compression),}$$

$T_2 = -650$ lb. (compression), $T_3 \cong 1440$ lb. (tension).

4.3 The diagonals of the rigid 6"x8" frame $P_1P_2P_3P_4$ are elastic cords tied together at their intersection Q. (Figure 4.7). A force $\vec{F}$ of magnitude one ounce is applied in the direction shown in the figure. Find the approximate displacement of Q in each of the following cases.

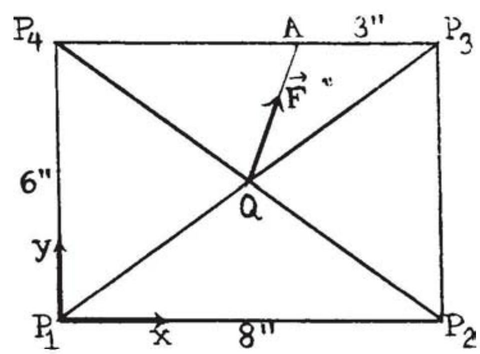


Figure 4.7

(a) The unstretched length of each of the two cords was 5 inches, and when stretched to the length of the diagonals the tension in each cord is 5 pounds.

(b) The unstretched length of each cord is 10 inches, and each will be stretched 2 inches by a one pound pull.

Answers. (a) $d = \begin{pmatrix} .008 \\ .042 \end{pmatrix}$ in.; (b) $d = \begin{pmatrix} .016 \\ .084 \end{pmatrix}$ in.

4.4 In Figure 4.3 let $P_1Q = P_2Q = L$ and let both members have the same elastic constant k. Show that the force F required to give a horizontal displacement d of Q is

$$F = \frac{k}{L^2} d^3 + \text{higher order terms.}$$

[Hint. Show that

$$F = 2kd \frac{\sqrt{L^2 + d^2} - L}{\sqrt{L^2 + d^2}} .$$

and expand in powers of d.]

5. Eigenvectors and Eigenvalues.

In Example 4.1 we saw how a force F applied at a point of a structure caused a displacement d of the point. It was noticeable that for the particular force F considered in the example the displacement d had a direction differing from that of F. In a problem of this sort it is obviously of some interest to inquire if the displacement can ever be in the same direction as the force. From the symmetry of this example it is evident that this will be so if the force is in the z-direction, but there may be (in fact, there are) other directions with the same property. It is not at all evident how one could answer the question in a general non-symmetrical case

Let us look at this question from the algebraic point of view. Corresponding to equation (4.5) we have a linear transformation

$$Av = w, \tag{5.1}$$

where v and w are n-component column vectors and A is an n x n matrix. Each v determines a w, and we want to find those particular v's, if any, whose corresponding w's have the same direction. If w has the same direction as v it must be of the form λv, where λ is a scalar. Thus we are interested in solutions of

$$Av = \lambda v. \tag{5.2}$$

We can generalize equations (5.1) and (5.2) by considering any vector space V and any linear transformation $Lv = w$ from V to V. Corresponding to (5.2) we then have the equation $Lv = \lambda v$. Equations of this form turn out to be of importance in a surprisingly wide spread of fields in both pure and applied mathematics. They have been studied extensively and have a rich theory with many ramifications. For most of this chapter we shall limit our considerations to the matrix case (5.2). For convenience we assume for the rest of this chapter that A is an n x n matrix.

Evidently $v = 0$ is a solution of (5.2) for any A and any λ. Excluding this trivial case, we can then state the general

Eigenvalue Problem as follows: Given a square matrix A, find all scalars λ and associated non-zero vectors v such that

$$Av = \lambda v \,. \tag{5.2}$$

such a scalar λ is called an eigenvalue* of A, and an associated vector v an eigenvector of A.

Example 5.1. Let

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad v = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Then

$$Av = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} -c_1 \\ 2c_2 \\ 0 \end{pmatrix}$$

so that if $Av = \lambda v$ we must have

$$\begin{pmatrix} -c_1 \\ 2c_2 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \quad .$$

That is,

$$-c_1 = \lambda c_1, \; 2c_2 = \lambda c_2, \; 0 = \lambda c_3.$$

There are three possibilities:

(1) $c_1 \neq 0$. $\lambda = -1$, $c_2 = 0$, $c_3 = 0$.

*The original English terms were "characteristic value" or "proper value." The German word is "Eigenwert." Somewhere along the line there was cross-breeding, and "eigenvalue" is now most common in this country.

$$\text{Eigenvalue} = -1, \text{ eigenvector} = \begin{pmatrix} c_1 \\ 0 \\ 0 \end{pmatrix} .$$

(2) $c_2 \neq 0$, $\lambda = 2$, $c_1 = 0$, $c_3 = 0$.

$$\text{Eigenvalue} = 2, \text{ eigenvector} = \begin{pmatrix} 0 \\ c_2 \\ 0 \end{pmatrix} .$$

(3) $c_3 \neq 0$, $\lambda = 0$, $c_1 = 0$, $c_2 = 0$.

$$\text{Eigenvalue} = 0, \text{ eigenvector} = \begin{pmatrix} 0 \\ 0 \\ c_3 \end{pmatrix} .$$

Note that zero eigenvalues can occur. We shall see later (cf. Problem 5.5) that only singular matrices have zero eigen- values.

Example 5.2. Let

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad v = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} .$$

As before, we get

$$-c_1 = \lambda c_1, \; 2c_2 = \lambda c_2, \; 2c_3 = \lambda c_3.$$

Here there are only two cases:

(1) As before.

(2) $c_1 = 0$ and either $c_2 \neq 0$ or $c_3 \neq 0$ (or both). We must have

$\lambda = 2$, and then there is no restriction on c_2 or c_3. Hence any non-zero vector of the form $\begin{pmatrix} 0 \\ c_2 \\ c_3 \end{pmatrix}$ is an eigenvector associated with the eigenvalue 2.

Note that a single eigenvalue may have more than one independent eigenvector associated with it. The situation in general is stated in the following theorem.

Theorem 5.1. If $v_1, v_2, \ldots, v_m$ are eigenvectors associated with the same eigenvalue λ then any non-zero linear combination of them is an eigenvector associated with λ.

Proof. If

$$Av_1 = \lambda v_1,\ Av_2 = \lambda v_2, \ldots, Av_m = \lambda v_m,$$

then

$$\begin{aligned} & A(c_1 v_1 + c_2 v_2 + \ldots + c_m v_m) \\ &= c_1 A v_1 + c_2 A v_2 + \ldots + c_m A v_m \\ &= c_1 \lambda v_1 + c_2 \lambda v_2 + \ldots + c_m \lambda v_m \\ &= \lambda (c_1 v_1 + c_2 v_2 + \ldots + c_m v_m). \end{aligned}$$

In particular we are free to change the length of an eigenvector by any desired non-zero factor, and we may make use of this freedom, if we wish, to normalize the vector. Note that there is never a unique eigenvector associated with a given eigenvalue.

Corollary 5.1. The set of all eigenvectors associated with a given eigenvalue, together with the zero vector, constitutes a linear subspace of V_n.

The eigenvectors associated with a given eigenvalue can therefore be specified by giving a basis of the corresponding subspace.

We now come to a basic theorem of the whole eigenvalue theory for matrices.

Theorem 5.2. The eigenvalues of A are the roots of the equation $f(\lambda) = 0$, where $f(\lambda)$ is the polynomial defined by

$$f(\lambda) = \det (\lambda I - A) .$$

Proof. Our basic equation (5.2) can be written in the form

$$\lambda v - Av = 0 ,$$

or, introducing the identity matrix I,

$$\lambda Iv - Av = 0 ,$$

or, finally

$$(\lambda I - A)v = 0 . \tag{5.3}$$

From the theory of Chapter 2, this system of n homogeneous equations in n unknowns has a non-trivial solution v if and only if the matrix of coefficients of v is singular, that is, if and only if $\det(\lambda I - A) = 0$. Hence the roots of this equation are precisely the eigenvalues of A. That $\det(\lambda I - A)$ is a polynomial in λ is evident from the definition of a determinant, which involves only multiplications and additions of the elements.

It is also evident that since λ appears only in the diagonal elements of the determinant the highest power of λ will arise from the product of these elements, and hence that $f(\lambda)$ has the form

$$f(\lambda) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \dots + c_n .$$

$f(\lambda)$ is called the <u>characteristic function</u> of A, and the equation $f(\lambda) = 0$ the <u>characteristic equation</u>.*

Since the characteristic equation is of degree n it has n roots, and so an n x n matrix has n eigenvalues. These eigenvalues may not be distinct, but may occur with certain multiplicities when considered as roots of the characteristic equation. (cf. Section 4 of Chapter 4, especially Corollary 4.3 and related material.) There is of course the possibility that some of the roots of $f(\lambda) = 0$ are complex numbers. Are we justified in calling them "eigenvalues," since according to our definition an eigenvalue is a scalar, i.e., a real number? In Section 10 we show how to handle this problem in a systematic

*One might think to call $f(\lambda)$ the "eigenfunction" but this word is reserved for a different use.

fashion, by redefining "scalar" to be a complex number.

Theorem 5.2 enables us in principle to solve the eigenvalue problem completely. We set up the characteristic function, solve the equation to get the eigenvalues λ, and then for each of the roots λ solve the system of equations (5.3) to get the associated eigenvectors v. The only trouble is that the amount of work involved when n is even as large as 4 or 5 is very great, and for larger values of n this direct approach is unsuitable. In a later section we shall consider some methods that are less direct but more suitable to machine programming. In the meantime, however, a few examples will illustrate how the direct approach works and what situations can arise.

Example 5.3. $A = \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}$.

$$f(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda-1 & -2 \\ -5 & \lambda-4 \end{vmatrix}$$

$$= (\lambda - 1)(\lambda - 4) - 10 = \lambda^2 - 5\lambda - 6.$$

Eigenvalues: $\lambda_1 = 6$, $\lambda_2 = -1$.

To get the eigenvectors solve $(\lambda I - A)v = 0$, where $v = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$.

(a) $\lambda_1 = 6$.

$$5c_1 - 2c_2 = 0,$$

$$-5c_1 + 2c_2 = 0.$$

$$5c_1 = 2c_2 .$$

Take $c_1 = 2$, $c_2 = 5$ for one independent solution, which is a basis of the 1-dimensional subspace of eigenvectors associated with the eigenvalue 6. We get, then, for one eigenvalue and associated eigenvector

$$\lambda_1 = 6, \qquad v_1 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}.$$

(b) $\lambda_2 = -1$.

$$-2c_1 - 2c_2 = 0,$$

$$-5c_1 - 5c_2 = 0.$$

$$c_1 + c_2 = 0.$$

We may take

$$c_1 = 1, \qquad c_2 = -1.$$

The second eigenvalue and eigenvector are then

$$\lambda_2 = -1, \qquad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Note in this example that in each part the two equations in the two variables c_1 and c_2 were linearly dependent, so that there was a non-trivial solution. Of course this must be so if the arithmetic is properly (and exactly) carried out, for the characteristic equation was constructed with precisely this end in view. This provides a check on the computation of the eigenvalues and the set-up of equations (5.3).

Example 5.4.

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} .$$

$$f(\lambda) = \begin{vmatrix} \lambda & -1 & -1 \\ 1 & \lambda-2 & -1 \\ 1 & -1 & \lambda-2 \end{vmatrix} = \lambda(\lambda-2)^2 + (-1)(-1)\cdot 1 + (-1)(-1)1$$

$$-\lambda(-1)(-1) - (\lambda-2)(-1)1 - (\lambda-2)(-1)1$$

$$= \lambda^3 - 4\lambda^2 + 5\lambda - 2.$$

Obviously $f(1) = 0$. Dividing $f(\lambda)$ by $\lambda - 1$ gives $\lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$. Hence the roots are $\lambda = 1,1,2$ and the distinct eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$.

(a) $\lambda_1 = 1$.

$$c_1 - c_2 - c_3 = 0,$$

$$c_1 - c_2 - c_3 = 0,$$

$$c_1 - c_2 - c_3 = 0.$$

We can choose two of the unknowns at will, so we have a 2-dimensional subspace of solutions. A possible basis is the pair of vectors

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

(b) $\lambda_2 = 2.$

$$2c_1 - c_2 - c_3 = 0,$$

$$c_1 - c_3 = 0,$$

$$c_1 - c_2 = 0.$$

It is obvious that the first equation is the sum of the other two, and that these two are independent, so there is no need to go through the formal reduction of these equations to echelon form. We can get a solution by assigning a value, 1 for instance, to any one of the variables. This gives the eigenvector

$$v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

You may have noted that the eigenvalue 1 is a two-fold root of the characteristic equation and the dimension of the corresponding linear subspace is 2. This is not just an accident, although these two numbers are not always equal, as Problem 5.2(b) shows. The exact relationship is important enough to warrant our stating it as a theorem.

Theorem 5.3. The dimension of the subspace of eigenvectors associated with a given eigenvalue never exceeds the multiplicity of that eigenvalue as a root of the characteristic equation.

For a proof of this theorem the reader is referred to

P. R. Halmos, Finite Dimensional Vector Spaces, Van Nostrand Co., Princeton, 1958.

We give one more example, to show how annoying the situation can get in even a simple case.

Example 5.5.

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} .$$

Here we find that $f(\lambda) = \lambda^3 - \lambda^2 - 1$. Not only are the roots of $f(\lambda) = 0$ irrational but two of them are complex. By Newton's Method (Thomas, Section 9-3), or some other method of approximation, we find that the real root is $\lambda = 1.466$ to 3 places. Equations (5.3) become

$$0.466c_1 - 2.000c_2 + 1.000c_3 = 0,$$

$$1.466c_2 - 1.000c_3 = 0,$$

$$-1.000c_1 - 1.000c_2 + 1.466c_3 = 0 .$$

Gauss elimination gives

1	-4.2918	2.1459
0	1.4660	-1.0000
0	-5.2918	3.6119
1	-4.2918	2.1459
0	1	-0.6821
0	0	0.0024

The occurrence of 0.0024 where there should be a zero is due to the error introduced by using the approximation 1.466 for the true (irrational) value of the eigenvalue. If we replace 0.0024 by zero we get a non-trivial solution,

$$c_1 = 0.781, \quad c_2 = 0.682, \quad c_3 = 1.$$

These values of c_1 and c_2 may well be incorrect by as much as .003. If we want 3-place accuracy in c_1 and c_2 we must determine λ to at least 4 places.

Beyond the added difficulty of computation we encounter no trouble in extending our methods to complex numbers.

<u>Example 5.6</u>.

$$A = \begin{pmatrix} -1 & 2 & 2 \\ -2 & 3 & 2 \\ 0 & -1 & 1 \end{pmatrix}.$$

Here $f(\lambda) = \lambda^3 - 3\lambda^2 + 5\lambda - 3 = (\lambda-1)(\lambda^2-2\lambda+3)$, and the eigenvalues are 1, $1 \pm \sqrt{2}\, i$. $\lambda_1 = 1$ gives the equations

$$2c_1 - 2c_2 - 2c_3 = 0,$$

$$2c_1 - 2c_2 - 2c_3 = 0,$$

$$c_2 \qquad = 0,$$

with the solution $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. For $\lambda_2 = 1 + \sqrt{2}\, i$ we get

$$
\begin{aligned}
(2+\sqrt{2}\, i)c_1 \qquad\qquad - 2c_2 - 2c_3 &= 0,\\
2c_1 + (-2+\sqrt{2}\, i)c_2 - 2c_3 &= 0,\\
c_2 + \sqrt{2}\, ic_3 &= 0.
\end{aligned}
$$

If we set $c_3 = 1$ the last equation gives $c_2 = -\sqrt{2}\, i$. Substituting these in the second equation gives

$$c_1 = (1 - \tfrac{1}{2}\sqrt{2}\, i)(-\sqrt{2}\, i) + 1 = -\sqrt{2}\, i - 1 + 1 = -\sqrt{2}\, i.$$

As a check we substitute in the first equation, to get

$$
\begin{aligned}
&(2+\sqrt{2}\, i)(-\sqrt{2}\, i) - 2(-\sqrt{2}\, i) - 2\\
&= -2\sqrt{2}\, i + 2 + 2\sqrt{2}\, i = 0.
\end{aligned}
$$

Hence $v_2 = \begin{pmatrix} -\sqrt{2}\, i \\ -\sqrt{2}\, i \\ 1 \end{pmatrix}$. Finally, since λ_3 is just the complex conjugate of λ_2 it is easy to see that the computation of v_3 will just change the sign of i throughout, and that therefore

$$v_3 = \begin{pmatrix} \sqrt{2}\, i \\ \sqrt{2}\, i \\ 1 \end{pmatrix}.$$

We shall see in the next chapter that matrices having complex eigenvalues are of frequent occurrence, and that computations of the above type are required in a wide variety of problems. However, there is one important class of matrices for which these complications never arise, namely the symmetric matrices. These matrices also have other convenient properties, some of which will be developed here and others deferred to later sections of this chapter.

First we need one observation about vectors with complex components. We have several times made use of the fact that $v^t v = v \cdot v \neq 0$ if $v \neq 0$. This is no longer true if we allow complex numbers, as is easily seen by the example

$$v = \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad v^t v = (1,i)\begin{pmatrix} 1 \\ i \end{pmatrix} = 1 + i^2 = 1 - 1 = 0.$$

To avoid this possibility we shall use $\bar{v}^t v$ instead of $v^t v$, where $\bar{v}$ is the complex conjugate of v, that is, the vector obtained by replacing each component of v by its conjugate. Thus

$$\text{if } v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \text{ then } \bar{v} = \begin{pmatrix} \bar{a}_1 \\ \vdots \\ \bar{a}_n \end{pmatrix} \text{ and}$$

$$\bar{v}^t v = \sum_{i=1}^{n} \bar{a}_i a_i = \sum_{i=1}^{n} |a_i|^2 \neq 0 \text{ if } v \neq 0.$$

This technique will be developed fully in Section 10.

We can now show that symmetric matrices have the desirable property mentioned above.

Theorem 5.4. The eigenvalues of a real symmetric matrix are real.

Proof. Let A be a symmetric matrix whose elements are real numbers. Then $A^t = A$ and $\overline{A} = A$. If λ is any eigenvalue of A and v an associated eigenvector we have $Av = \lambda v$. Multiplying on the left by $\overline{v}^t$ gives

$$\overline{v}^t Av = \lambda \overline{v}^t v. \tag{5.4}$$

Now take the transpose of both sides of this equation and then the conjugate. Using the properties of A we get in succession

$$\begin{aligned} v^t A\overline{v} &= \lambda v^t \overline{v}, \\ \overline{v}^t Av &= \overline{\lambda}\overline{v}^t v . \end{aligned} \tag{5.5}$$

Subtracting (5.5) from (5.4) gives

$$0 = (\lambda - \overline{\lambda})\overline{v}^t v .$$

Since v is an eigenvector it is not zero, and hence, as we have seen, $\overline{v}^t v \neq 0$. It follows that $\lambda = \overline{\lambda}$, and so λ is a real number (Chapter 4, Problem 2.15(h)).

In much the same way we can prove another useful property of symmetric matrices.

Theorem 5.5. Eigenvectors of a symmetric matrix that are associated with different eigenvalues are orthogonal.

Proof. Let $A^t = A$, let $\lambda \neq \mu$, and let

$$Av = \lambda v, \tag{5.6}$$

$$Aw = \mu w. \tag{5.7}$$

Multiply (5.6) on the left by w^t and (5.7) on the left by v^t; this gives

(5.8) $$w^t A v = \lambda w^t v,$$

(5.9) $$v^t A w = \mu v^t w \; .$$

If we take the transpose of (5.8) we get

$$v^t A w = \lambda v^t w,$$

and subtracting this from (5.9) gives

$$0 = (\mu - \lambda) v^t w \; .$$

Since $\mu \neq \lambda$ we must have $v^t w = 0$, that is, v and w are orthogonal.

From this theorem we can see that the eigenvectors of a symmetric matrix have an interesting geometric structure. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the distinct eigenvalues of a symmetric matrix A, λ_i being a root of multiplicity s_i of the characteristic equation. By Theorem 5.4 the λ_i are all real, and $s_1 + s_2 + \ldots + s_m = n$, the degree of the equation. With each λ_i there is associated a vector subspace W_{r_i} of V_m consisting of all the eigenvectors of A associated with λ_i. By Theorem 5.3, $r_i \leqq s_i$. By Theorem 5.5 any vector in a W_{r_i} is orthogonal to any vector in a W_{r_j} if $i \neq j$. Now each W_{r_i}, being r_i-dimensional, has a basis of r_i elements, and by the Schmidt Process, we can orthogonalize this basis. We therefore get altogether a set of vectors of the following sort:

$v_1, v_2, \ldots, v_{r_1}$, eigenvectors of λ_1;

$v_{r_1+1}, \ldots, v_{r_1+r_2}$, eigenvectors of λ_2;

. . . .

$v_{r_1+\ldots+r_{m-1}+1}, \ldots, v_{r_1+\ldots+r_m}$, eigenvectors of λ_m.

By the above remarks all $r_1 + \ldots + r_m$ of these vectors are orthogonal, and they form an orthogonal basis for the $(r_1 + \ldots + r_m)$-dimensional subspace spanned by <u>all</u> the eigenvectors of A. It is of great importance that $r_1 + \ldots + r_m = n$; that is:

<u>Theorem 5.6</u>. The eigenvectors of a symmetric n x n matrix span V_n.

The proof of this theorem is more elaborate than we wish to give in this course; the interested reader can look it up in Halmos, loc. cit.

From the facts that $\sum r_i = n$, $\sum s_i = n$, and $r_i \leqq s_i$, we at once conclude that $r_i = s_i$. That is:

<u>Corollary 5.2</u>. For a symmetric matrix the dimension of the subspace of eigenvectors associated with an eigenvalue is equal to the multiplicity of this eigenvalue.

<u>Example 5.7</u>.

$$A = \begin{pmatrix} -2 & 2 & 2 \\ 2 & 1 & 4 \\ 2 & 4 & 1 \end{pmatrix} .$$

We find $f(\lambda) = \lambda^3 - 27\lambda - 54$, and the eigenvalues are 6, -3, -3.

(a) $\lambda_1 = 6$.

$$8c_1 - 2c_2 - 2c_3 = 0,$$

$$-2c_1 + 5c_2 - 4c_3 = 0,$$

$$-2c_1 - 4c_2 + 5c_3 = 0.$$

These reduce to the echelon form

$$\begin{matrix} 1 & 0 & -1/2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{matrix}$$

giving a solution

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

(b) $\lambda_2 = -3$.

$$-c_1 - 2c_2 - 2c_3 = 0,$$

$$-2c_1 - 4c_2 - 4c_3 = 0,$$

$$-2c_1 - 4c_2 - 4c_3 = 0.$$

Here we have two independent solutions, taking, for example,

$v_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$. However, these are not orthogonal,

so we proceed to orthogonalize them. (For convenience we write them as row vectors.)

$$v_2' = v_2.$$

$$v_3' = v_3 - \frac{v_3 \cdot v_2'}{v_2' \cdot v_2'} v_2'$$

$$= (-2,0,1) - 1/2(0,-1,1) = 1/2(-4,1,1).$$

Hence we can take as our final orthogonal basis of eigenvectors of A the set

$(1,2,2)$, eigenvalue 6;

$(0,-1,1)$, $(-4,1,1)$, eigenvalue 3.

As an application of these results we return to the considerations of Section 4. The matrix M is defined by $M = UKU^t$, and so

$$M^t = (UKU^t)^t = (U^t)^t K^t U^t = UKU^t = M,$$

that is, M is symmetric, and so is $A = M^{-1}$. We have therefore exactly three possible cases.

(1) Three distinct eigenvalues. There are exactly three lines through Q, mutually perpendicular, along which F may act to insure that d will be in the same direction. The stiffness ratio $\|F\|/\|d\|$ will be different in each of these directions (being just the associated eigenvalue of M).

(2) Two distinct eigenvalues. There is a line through Q, and a plane through Q perpendicular to this line, such that if F acts either along this line or along any line in the plane then d will be in the same direction as F. The ratio $\|d\|/\|F\|$ will be the same for all directions in the plane but

different for the perpendicular line.

(3) One distinct eigenvalue λ. $d = \lambda F$ for all directions of F.

Problems

5.1 Problem 4.1 leads to the matrix

$$A = \begin{pmatrix} 1.96 & 0.71 \\ 0.71 & 2.72 \end{pmatrix} \times 10^{-5}$$

transforming F into d. Find the eigenvectors of A and draw them on a picture of the structure. Do they seem physically reasonable?

Answer: $\begin{pmatrix} 1 \\ 1.67 \end{pmatrix}$, $\begin{pmatrix} 1.67 \\ -1 \end{pmatrix}$.

5.2 Find the eigenvalues and eigenvectors of

(a) $\begin{pmatrix} 4 & 12 \\ 3 & 4 \end{pmatrix}$, (b) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$, (c) $\begin{pmatrix} 0 & -2 & 2 \\ 2 & -1 & 0 \\ 2 & -2 & 1 \end{pmatrix}$,

(d) $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$, (e) $\begin{pmatrix} 0 & 0 & -1 & 0 & 1 \\ -2 & 1 & 1 & 2 & -1 \\ 2 & -2 & 0 & -2 & -1 \\ 1 & 0 & -1 & -1 & 1 \\ 2 & -2 & -1 & -2 & 0 \end{pmatrix}$,

(f) $\begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix}$, (g) $\begin{pmatrix} 3 & 4 & 2 \\ -2 & -2 & -1 \\ 0 & -1 & 0 \end{pmatrix}$, (h) $\begin{pmatrix} 1 & -3 & 1 & 1 \\ -3 & 1 & 1 & 1 \\ 1 & 1 & 1 & -3 \\ 1 & 1 & -3 & 1 \end{pmatrix}$.

(i) $\begin{pmatrix} -1 & -2 & -2 \\ 2 & 3 & 2 \\ -2 & -2 & -1 \end{pmatrix}$

Answers: (a) 10,(2,1); -2,(2,-1).

(b) 1,(1,1,0); 2,(1,1,1).

(c) 0,(1,2,2); 1,(2,2,3); -1,(0,1,1).

(d) 2±1.618,(1,±1.618,1.618,±1);2±0.618,

(1,±.618,-.618,∓1).

[Note. In expanding det(λI-A) let λ-2 = μ.]

(e) 0,(1,0,0,1,0); 1,(0,1,-1,0,1); -1,

(0,-1,1,1,1),(0,0,1,0,1).

(f) 2±i, (1,1±i).

(g) 1,(1,-1,1); ±i, (∓i-1,1,±i).

(h) 0,(1,1,1,1); -4,(1,1,-1,-1); 4,(1,-1,0,0),

(0,0,1,-1).

(i) -1,(1,-1,1); 1,(1,-1,0), (1,0,-1).

5.3 Find the eigenvalues and eigenvectors of

$$(a)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (b)\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (c)\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

(d) Generalize (a), (b), (c) to n x n matrices.

5.4 (a) The symmetric matrix

$$A = \begin{pmatrix} -1 & -1 & -1 & 1 \\ -1 & -1 & 1 & -1 \\ -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$$

has distinct eigenvalues 2 and -2. Find an orthogonal basis of eigenvectors and the multiplicity of each eigenvalue. [It is not necessary to find the characteristic function.]

(b) Do the same for the matrix

$$B = \begin{pmatrix} 0 & 3 & 6 & 6 \\ 3 & 8 & -2 & -2 \\ 6 & -2 & -4 & 5 \\ 6 & -2 & 5 & -4 \end{pmatrix}$$

whose distinct eigenvalues are 9 and -9.

5.5 $\lambda = 0$ is an eigenvalue of A if and only if $Av = 0$ has a non-zero solution. Hence show that only singular matrices have zero eigenvalues.

5.6 Prove the following facts directly from the definitions of eigenvalue and eigenvector.

(a) If λ is an eigenvalue of A and v an associated eigenvector, then

(i) For any scalar c, $c\lambda$ is an eigenvalue of cA, and $\lambda+c$ of $A + cI$, each with associated eigenvector v.

(ii) λ^2 is an eigen~~vector~~ of A^2, λ^3 of A^3, etc., each with associated eigenvector v.

(iii) If A is non-singular, then λ^{-1} is an eigenvalue of A^{-1} with associated eigenvector v.

(b) If A and B have respective eigenvalues λ and μ, with a common associated eigenvector v, then $\lambda+\mu$ is an eigenvalue of A+B, and $\lambda\mu$ of AB, each with associated eigenvector v.

5.7 Let $A = (a_{ij})$ have the characteristic function

(5.10) $f(\lambda) = \det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + \dots + c_n$.

If $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A, not necessarily distinct, we have (Corollary 4.3 of Chapter 4)

$$f(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)\dots(\lambda-\lambda_n) .$$

(a) Show that $\sum_{i=1}^{n} \lambda_i = -c_1$ and $\prod_{i=1}^{n} \lambda_i = (-1)^n c_n$.*

(This is a well-known theorem in algebra.)

(b) Show that $\det A = (-1)^n c_n = \prod_{i=1}^{n} \lambda_i$.

[Put $\lambda = 0$ in (5.10).]

(c) Show that

$$\det(\lambda I - A) = (\lambda - a_{11})(\lambda - a_{22})\dots(\lambda - a_{nn})$$

+ terms of degree at most n-2,

and that therefore $\sum_{i=1}^{n} a_{ii} = -c_1 = \sum_{i=1}^{n} \lambda_i$.

The quantity $\sum_{i=1}^{n} a_{ii}$ is called the <u>trace</u> of A,

abbreviated tr(A).

5.8. The matrix $\begin{pmatrix} 0 & -2 & 2 \\ 2 & -1 & 0 \\ 2 & -2 & k \end{pmatrix}$ has eigenvalues 0,1,-1.

What is the value of k?

---

* $\prod_{i=1}^{n} \lambda_i$ means the product $\lambda_1\lambda_2\cdots\lambda_n$.

6. Expansion in Eigenvectors. Numerical Computation.

The importance of Theorem 5.6 resides in the fact that the eigenvectors of a given symmetric matrix A can be chosen as an orthogonal (or orthonormal) basis of V_n, and in terms of this basis many properties of A can be expressed quite simply. We shall consider here only a few such properties.

Let $v_1, v_2, \ldots, v_n$ be an orthogonal basis of eigenvectors of A, associated with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively, that is

(6.1) $$Av_i = \lambda_i v_i, \quad i = 1, \ldots, n.$$

Note that the λ_i need not be all different from one another.

Since $v_1, \ldots, v_n$ is a basis, any vector w can be expressed in the form

(6.2) $$w = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = \sum_{i=1}^{n} a_i v_i .$$

Since the v_i are pairwise orthogonal Theorem 3.3 can be used to compute the coefficients a_i, thus:

(6.3) $$a_i = \frac{w \cdot v_i}{v_i \cdot v_i}, \quad i = 1, \ldots, n.$$

(If the v's are normalized the denominator in (6.3) is 1.) From (6.2) we get

$$Aw = A \sum a_i v_i = \sum a_i A v_i ,$$

and so from (6.1)

$$\text{(6.4)} \qquad Aw = \sum a_i \lambda_i v_i .$$

In Chapter 8 equations (6.2) and (6.3), applied to vector subspaces of C^∞, will be used to develop the theory of Fourier Series. (See also Problem 10.5.) For the present we make no use of (6.3). However, the passage from (6.2) to (6.4) has an interesting geometrical interpretation. If the v's are normalized, the scalars $a_1,\ldots,a_n$ appearing in (6.2) are just the components of w with respect to the orthonormal basis of the v's (cf. Section 3). Equation (6.4) tells us that referred to this basis multiplication by A just multiplies each component by the corresponding eigenvalue. Thus the transformation associated with a symmetric matrix has a simple structure, being a combination of stretchings or shrinkings in a set of properly chosen orthogonal directions. We shall return to this viewpoint in the next section.

Equation (6.4) is also the basis of an important method of computing eigenvalues and eigenvectors. As was remarked in the previous section the direct methods used there are of little value if n is large. The method we are about to give is not of completely general application but it is easily programmed and in many types of problems gives all the needed information about the matrix.

Let A be a symmetric matrix, with the additional property of having a dominant eigenvalue, that is, one which is greater in absolute value than any other eigenvalue. We may call this eigenvalue λ_1, so that we have

(6.5) $$|\lambda_1|>|\lambda_i|, \quad i = 2,3,\dots,n.$$

By a repetition of the process we used in passing from (6.2) to (6.4) we can get

(6.6) $$A^k w = a_1\lambda_1^k v_1 + a_2\lambda_2^k v_2 + \dots + a_n\lambda_n^k v_n.$$

Let us assume that w is so chosen that $a_1 \neq 0$. Then (6.6) can be written

(6.7) $$A^k w = a_1\lambda_1^k \left[v_1 + \frac{a_2}{a_1}\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \dots + \frac{a_n}{a_1}\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n \right].$$

From (6.5), $|\lambda_2/\lambda_1| < 1$, and so for large enough k, $(\lambda_2/\lambda_1)^k$ can be made as small as we please. Similarly all the terms in the brackets beyond the first can be made very small compared with the first term. Hence for large k

(6.8) $$A^k w \cong a_1\lambda_1^k v_1.$$

If we designate the vector $A^k w$ by w_k, and write two successive steps of the form (6.8)

$$w_k \cong a_1\lambda_1^k v_1,$$

$$w_{k+1} \cong a_1\lambda_1^{k+1} v_1,$$

we conclude that

(6.9) $$w_{k+1} \cong \lambda_1 w_k.$$

This approximation, together with the recursion formula for w_k,

(6.10) $$w_{k+1} = Aw_k$$

gives us a recursive method for determining λ_1. Furthermore, w_k is an approximation to an associated eigenvector.

Example 6.1. Let

$$A = \begin{pmatrix} 2.8 & -0.4 \\ -0.4 & 2.2 \end{pmatrix},$$

for which $\lambda_1 = 3$, $\lambda_2 = 2$, and we can take as eigenvectors

$$v_1 = \begin{pmatrix} 1 \\ -.5 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

(Figure 6.1). If we take

$$w = v_1 + v_2 = \begin{pmatrix} 2 \\ 1.5 \end{pmatrix}$$

then

$$w_1 = Aw = \begin{pmatrix} 5 \\ 2.5 \end{pmatrix}$$

$$= 3v_1 + 2v_2.$$

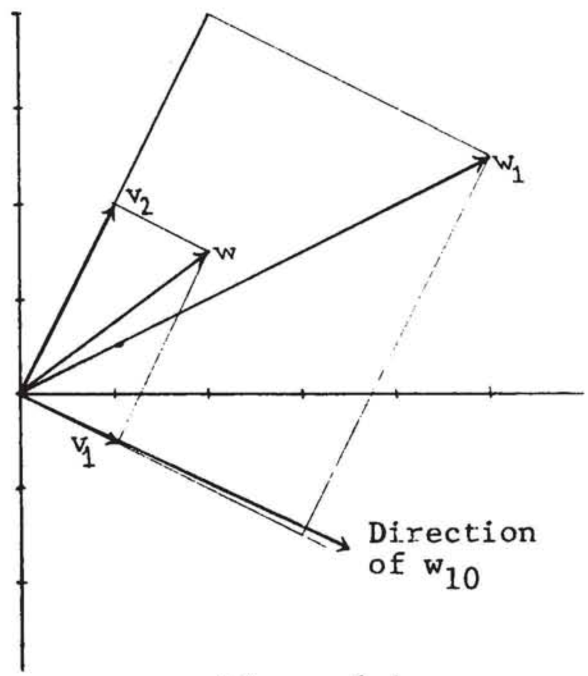


Figure 6.1

It is evident from the figure that w_1 tends more to the direction of v_1 than does w. If we continue the process for nine more steps we get

$$w_{10} = 3^{10}v_1 + 2^{10}v_2 = 3^{10}(v_1 + .017v_2),$$

whose direction, indicated in the figure, is very nearly that of the dominant eigenvector v_1.

(Another geometrical interpretation of this process is given in Section 9.)

In this example we started with known eigenvalues, to see the geometric meaning of the process. The next example shows how the process can be used to find the dominant eigenvalue and associated eigenvector.

Example 6.2. Assuming that

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

has a dominant eigenvalue, find it.

We pick $w_0 = w$ more or less at random. For hand computation a good choice is a column vector of dominant length, if there is one. Writing the w's as row vectors for convenience we get by successive multiplication by A:

k	w_{k1}	w_{k2}	w_{k3}
0	2	2	1
1	9	7	4
2	36	29	16
3	146	117	65
4	591	474	263
5	2393	1919	1065
6	9689	7770	4312

$w_5 = 2393(1,0.80192,0.44505)$.

$w_6 = 9689(1,0.80194,0.44504)$.

Hence $\lambda_1 = 9689/2393 = 4.049$. Also, $v_1 = (1,0.8019,0.4450)$ is the associated eigenvector.

Comment 1. Since we are interested only in the ratios cf successive w's we can divide by any scalar at any stage of the process. This device can be used to keep the numbers within bound and to enable us to tell when we have reached the desired accuracy. Suppose, for example, we want λ_1 correct to two decimal places. We proceed as above, but at each step divide w_k by a scalar so as to reduce w_{k1} to 1. This gives:

k	w_{k1}	w_{k2}	w_{k3}
0	2	2	1
	1	1	0.5
1	4.5	3.5	2.0
	1	.778	.444
2	4.000	3.222	1.778
	1	.806	.444
3	4.056	3.250	1.806
	1	.801	.445
4	4.047	3.246	1.801
	1	.802	.445

Evidently we can stop here, with $\lambda_1 = 4.05$, $v = (1,.80,.45)$.

Comment 2. For the critical equation (6.7) to hold,it was necessary that the v_1-component of w, a_1, be different from zero. However even if we were so unfortunate as to pick a w_0 for which $a_1 = 0$ the round-off errors in computations like the above would be almost certain to introduce such a component, and the computation, after a slow start, would eventually converge to the desired value.

Comment 3. The rate of convergence of the computation is determined by the maximum of $|\lambda_i/\lambda_1|$ for $i > 1$. Convergence is essentially like that of a geometric series with this as the ratio. The fact that our example gave an accuracy of about .0005 in six steps indicates that, to an order of magnitude

$$\left|\frac{\lambda_2}{\lambda_1}\right|^6 \cong .0005, \quad \text{or } |\lambda_2| \cong 1,$$

if λ_2 is the eigenvalue of next largest absolute value.

Even if this method of repeated multiplication gave only the dominant eigenvalue it would be of use, for in many types of problems (vibration problems, for example) only the dominant eigenvalue is needed. However, having found λ_1 and v_1 we can easily modify the matrix A so as essentially to remove λ_1 and leave the remaining eigenvalues for independent consideration.

Theorem 6.1. Let $v_1, v_2, \ldots, v_n$ be an orthogonal set of eigenvectors of a symmetric matrix A, with associated eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. In particular let $\|v_1\| = 1$. Define the matrix B by $B = A - \lambda_1 v_1 v_1^t$. Then B is symmetric and has the eigenvectors $v_1, v_2, \ldots, v_n$ with the associated eigenvalues $0, \lambda_2, \ldots, \lambda_n$.

Proof.

(i) $$B^t = (A - \lambda_1 v_1 v_1^t)^t = A^t - \lambda_1 (v_1 v_1^t)^t$$

$$= A - \lambda_1 v_1 v_1^t = B,$$

so B is symmetric.

(ii) $$Bv_1 = (A - \lambda_1 v_1 v_1^t) v_1$$

$$= Av_1 - \lambda_1 (v_1 v_1^t) v_1$$

$$= \lambda_1 v_1 - \lambda_1 v_1 (v_1^t v_1)$$

$$= \lambda_1 v_1 - \lambda_1 v_1$$

$$= 0 v_1,$$

since $v_1^t v_1 = \|v_1\|^2 = 1$. Hence 0 is an eigenvalue of B with associated eigenvector v_1.

(iii) $$Bv_2 = (A - \lambda_1 v_1 v_1^t) v_2$$

$$= Av_2 - \lambda_1 (v_1 v_1^t) v_2$$

$$= \lambda_2 v_2 - \lambda_1 v_1 (v_1^t v_2)$$

$$= \lambda_2 v_2,$$

since $v_1^t v_2 = 0$ by orthogonality of the eigenvectors. Hence v_2 is an eigenvector of B with associated eigenvalue λ_2, and the same argument will obviously hold for $v_3, \ldots, v_n$.

If our given v_1 is not normalized we can either normalize it, that is replace it by $v_1/\|v_1\|$, or modify the definition of B by taking

$$B = A - \frac{\lambda_1}{v_1{}^t v_1} v_1 v_1{}^t \ .$$

<u>Example 6.2</u>, continued. We have

$v_1 = (1, 0.8019, 0.4450)$, $\lambda_1 = 4.049$, $\lambda_1/v_1{}^t v_1 = 2.199$.

Hence

$$B = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} - 2.199 \begin{pmatrix} 1 \\ 0.8019 \\ 0.4450 \end{pmatrix} (1,\ 0.8019,\ 0.4450)$$

$$= \begin{pmatrix} 2 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} - 2.199 \begin{pmatrix} 1 & 0.8019 & 0.4450 \\ 0.8019 & 0.6430 & 0.3568 \\ 0.4450 & 0.3568 & 0.1980 \end{pmatrix}$$

$$= \begin{pmatrix} -.199 & .237 & .021 \\ .237 & -.414 & .215 \\ .021 & .215 & -.435 \end{pmatrix} \ .$$

To find the dominant eigenvalue of this matrix we take the middle column as w_o.

k	w_{k1}	w_{k2}	w_{k3}
0	.237	- .414	.215
	1	-1.747	0.907
1	- .593	1.155	-0.748
	1	-1.948	1.261
2	- .633	1.315	-0.945
	1	-2.077	1.493
3	- .658	1.418	-1.074
	1	-2.155	1.632
4	- .674	1.480	-1.151
	1	-2.196	1.708

It is evident that the process is converging slowly, indicating that the two remaining eigenvalues do not differ much in absolute value. Completion of the computation (by machine!) gives $\lambda_2 = -.692$. Finally, we find $\lambda_3 = -.357$; since $|\lambda_3/\lambda_2| > .5$ the slowness of the convergence is explained.

The repeated multiplication process is not affected if the dominant eigenvalue is multiple, but it does break down if there is no dominant eigenvalue, that is, if $\lambda_2 = -\lambda_1$ and all other eigenvalues are of smaller absolute value. Even if λ_1 is dominant the process can converge too slowly to be useful, when $|\lambda_2/\lambda_1|$ is very close to 1. There are ways of generalizing the method to cover such cases, but on the whole they are complicated and difficult to apply. A good reference for various methods and their generalization is L. Fox, An Introduction to Numerical Linear Algebra, Oxford University Press, New York, 1965, Chapter 9.

Some of these methods can also be extended to non-symmetric matrices. In particular the repeated multiplication method will work equally well for any matrix that has a dominant eigenvalue.

Problems

6.1 Find to one decimal place the dominant eigenvalue and associated eigenvector of each of the following matrices, using the method of repeated multiplication:

(a) $\begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$, (b) $\begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, (c) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & -1 & 0 \end{pmatrix}$.

Answers: 3.6, (1.0, 1.3); 1.5, (1.2,1.0,1.5);
2.5, (1.0, -2.9, 1.5).

6.2 The eigenvalues of $\begin{pmatrix} 1 & -2 & 2 \\ -2 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix}$ are integers. Find them by the repeated multiplication process and use of Theorem 6.1.

6.3 The dominant eigenvalue of the matrix in Problem 6.1(c) is $\lambda_1 = 2.532$. Using the results of Problem 5.7 show that λ_2 and λ_3 satisfy the equations

$$\lambda_2+\lambda_3 = 0.468, \quad \lambda_2\lambda_3 = -1.185$$

and solve these for λ_2 and λ_3 accurate to three decimal places.

Answer: 1.348, - 0.880

6.4 From equation (6.4) it follows that if a symmetric matrix C has all its eigenvalues equal to zero then Cw = 0 for every vector w, and hence C = 0. Using this and repeated application of Theorem 6.1 show that any symmetric matrix A can be expressed in the form

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^t$$

where $v_1,\ldots,v_n$ are an orthonormal basis of eigenvectors and $\lambda_1,\ldots,\lambda_n$ the associated eigenvalues.

6.5 Using the notation and assumptions at the beginning of this section, show that

$$w^t A w = \sum_{i=1}^{n} \lambda_i (a_i \|v_i\|)^2.$$

In particular, if the v_i are orthonormal,

$$w^t A w = \sum_{i=1}^{n} \lambda_i a_i^2.$$

6.6 If the n x n matrix A has n independent eigenvectors $v_1, \ldots, v_n$ with the associated eigenvalues $\lambda_1, \ldots, \lambda_n$, show that

(a) The matrix $A - \lambda_1 I$ has eigenvectors $v_1, \ldots, v_n$ with the associated eigenvalues $0, \lambda_2 - \lambda_1, \lambda_3 - \lambda_1, \ldots, \lambda_n - \lambda_1$;

(b) The matrix $(A-\lambda_1 I)(A-\lambda_2 I)$ has eigenvectors $v_1, \ldots, v_n$ with the associated eigenvalues $0, 0, (\lambda_3-\lambda_1)(\lambda_3-\lambda_2), \ldots, (\lambda_n-\lambda_1)(\lambda_n-\lambda_2)$.

(c) By continuing in this way, the matrix $(A-\lambda_1 I)(A-\lambda_2 I)\ldots(A-\lambda_n I)$ is the zero matrix.

(d) Do you see any analogy between (c) and the characteristic equation $f(x) = 0$ of A?

6.7. Given

$$A = \begin{pmatrix} -1 & 1 & 1 \\ -1 & 0 & -1 \\ 2 & -1 & 0 \end{pmatrix} \quad ;$$

(a) Find the eigenvalues of A exactly, by solving the characteristic equation.

(b) Starting with $w_o = (-1,-1,2)$, carry out five steps of the repeated multiplication process to get estimates for an eigenvalue and its eigenvector. In what way do your estimates agree with the results of (a)?

6.8. The dominant eigenvalue of

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

is 2, and an associated eigenvector is $(1,0,-1)$. Apply Theorem 6.1 and use three steps of the repeated multiplication process to obtain an estimate of the next eigenvalue and its eigenvector.

6.9. Write a program to compute the dominant eigenvalue and eigenvector of a matrix A (assuming that A has a dominant eigenvalue.

Note. The method described and illustrated above for finding eigenvalues and eigenvectors is readily programmed for machine computation, with one modification. The device used in Comment 1 on Example 6.2 is not convenient for machine use since the component one chooses to divide by may turn out to approach zero. Instead of using this method to compare successive values of w it is more convenient, though involving slightly more arithmetic, to normalize w at each step. We can then compare successive values of normalized w, component by component, and stop the process when the maximum change for any component is sufficiently small.

Thus the repetitive process goes as follows:

$$u = Aw$$

$$c = 1 / \sqrt{u \cdot u}$$

$$z = cu$$

$$a = \max_{i=1,\cdots,n} |w_i - z_i|$$

$$b = \max_{i=1,\cdots,n} |w_i + z_i|$$

$$d = \min(a,b)$$

$$w = z$$

and repeat until d is less than some prescribed ξ.

[Note. If $\lambda > 0$, z becomes nearly equal to w, and a becomes small. If $\lambda < 0$, z becomes nearly equal to $-w$, and b becomes small. In either case d becomes small.]

An approximate value of λ is then obtainable by computing $w \cdot Aw$. For since w is an approximate eigenvector and $w \cdot w = 1$ we have $w \cdot Aw \cong w \cdot \lambda w = \lambda w \cdot w = \lambda$.

6.10. Extend the program of the previous problem to compute all the eigenvalues of a symmetric matrix, using Theorem 6.1.

7. Change of Basis.

Let us consider, briefly, the abstract point of view mentioned near the beginning of Section 5. If V is any vector space and L any linear transformation from V to V we can consider the problem of finding the eigenvalues λ and eigenvectors $\vec{v}$ that satisfy

$$L\vec{v} = \lambda\vec{v}. \tag{7.1}$$

From this statement of the problem it is evident that the eigenvalues and eigenvectors of a linear transformation have a meaning independent of any algebraic devices used for computing them.

If V has finite dimension n we can introduce a set of n basis vectors, in terms of which $\vec{v}$ is represented as a column vector v and L as a matrix A,* so that the abstract equation (7.1) assumes the familiar form

$$Av = \lambda v. \tag{7.2}$$

If a different basis for V were chosen the same abstract vector $\vec{v}$ would be represented by a different vector v' and the transformation L by a different matrix A'. Our argument in the previous paragraph then assures us that

$$A'v' = \lambda v' . \tag{7.3}$$

It is instructive to relate this situation to the material

*More formally, we set up an isomorphism between V and V_n in which $\vec{v} \longleftrightarrow v$. (See Theorem 8.2 of Chapter 2.)

on change of basis developed in Section 15 of Chapter 2. There we found that the relation between two bases in the same space V can be described by a non-singular matrix C, such that for two representatives v and v' of the same abstract vector $\vec{v}$ we have

$$v' = Cv. \tag{7.4}$$

Also, the two representative matrices A and A' of the same linear transformation L are related by

$$A' = CAC^{-1}. \tag{7.5}$$

From these relations we can give an independent derivation of (7.3) from (7.2), as follows:

$$A'v' = CAC^{-1}Cv = CAv = C\lambda v = \lambda Cv = \lambda v' .$$

This algebraic argument merely confirmed what we already knew, but similar arguments can give us useful additional information that might be very hard to get from a purely abstract approach. One such bit of information is the following.

Theorem 7.1. If C is non-singular then A and CAC^{-1} have the same characteristic equation.

Proof. First we note that

$$\lambda I-CAC^{-1} = C(\lambda I)C^{-1}-CAC^{-1} = C(\lambda I-A)C^{-1}.$$

Hence

$$\begin{aligned}\det(\lambda I-CAC^{-1}) &= (\det C)[\det(\lambda I-A)](\det C^{-1})\\ &= (\det C)[\det(\lambda I-A)](\det C)^{-1}\\ &= \det(\lambda I-A).\end{aligned}$$

This theorem tells us that not only the values of the λ's but also their multiplicities are independent of the choice of basis.

Example 7.1. Making up an example.

We want a simple, but non-trivial, matrix with eigenvalues 1, 1, 2. The simplest is obviously

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} .$$

To make this more interesting, without changing the eigenvalues, we replace it by CAC^{-1}, choosing C as some matrix with $\det C = 1$ to avoid fractions. (cf. Problem 14.7 of Chapter 2.) An easy way to do this is to start with C = I and apply elementary row and column operations of Type (ii), (Definition 11.1 of Chapter 2). With

$$C = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix} \text{ we get } CAC^{-1} = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} .$$

This is how Example 5.4 was constructed.

In this example we started with a diagonal matrix, for which the diagonal elements are obviously the eigenvalues, and by a change of basis put it in a more complicated form. We shall now show that the reverse process can in some cases be carried out, thus giving us at least a partial answer to a question raised on page 2.163, namely Question B: Can we choose a basis for which a given matrix A has a particularly simple form? The partial answer is the following.

Theorem 7.2. If A has n independent eigenvectors, and if these eigenvectors are taken as a new basis, then the corresponding matrix A' will be a diagonal matrix with the eigenvalues as the diagonal elements.

Proof. If the eigenvectors are taken to be the basis, then with respect to themselves as basis they are represented by the column vectors

$$v_1' = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad v_2' = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \text{etc.}$$

If $A' = (a_{ij}')$ we find that

$$A'v_1' = \begin{pmatrix} a_{11}' \\ a_{21}' \\ \vdots \\ a_{n1}' \end{pmatrix} = \lambda_1 v_1' = \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

and so $a_{11}' = \lambda_1$, $a_{21}' = \ldots = a_{n1}' = 0$. A similar examination of the other v_i' gives

$$A' = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} .$$

If we want the matrix C that effects the change of basis, this also is easily obtainable. If $v_1, \ldots, v_n$ are the eigenvectors of A, then C has the property that $v_i' = Cv_i$, or $v_i = C^{-1}v_i'$. By an argument similar to the one above, this last equation tells us that v_i is the i-th column vector of C^{-1}. Thus C^{-1}, and hence C, is determined.

We have seen (Theorem 5.6) that a symmetric matrix has n independent eigenvectors, and so every symmetric matrix is diagonalizable. The following theorem, often useful in itself, enables us to identify another class of diagonalizable matrices.

Theorem 7.3. If $v_1,\dots,v_m$ are eigenvectors of A whose eigenvalues $\lambda_1,\dots,\lambda_m$ are distinct, then $v_1,\dots,v_m$ are independent.

Proof. Suppose that $v_1,\dots,v_m$ are dependent. Then (Corollary 6.1 of Chapter 2) one of them, say v_k, is a combination of the previous ones $v_1,\dots,v_{k-1}$,

(7.6) $$v_k = a_1v_1 + \dots + a_{k-1}v_{k-1} ,$$

We may take k to be the smallest subscript with this property, so that $v_1,\dots,v_{k-1}$ are independent. If we multiply both sides of (7.6) by A we get

(7.7) $$\lambda_kv_k = \lambda_1a_1v_1 + \dots + \lambda_{k-1}a_{k-1}v_{k-1} .$$

Multiplying (7.6) by λ_k and subtracting from (7.7) gives

$$0 = (\lambda_1-\lambda_k)a_1v_1 + (\lambda_2-\lambda_k)a_2v_2 + \dots + (\lambda_{k-1}-\lambda_k)a_{k-1}v_{k-1}.$$

Since $v_1,\dots,v_{k-1}$ are independent vectors, this relation implies that

$$(\lambda_1-\lambda_k)a_1 = (\lambda_2-\lambda_k)a_2 = \dots = (\lambda_{k-1}-\lambda_k)a_{k-1} = 0,$$

and since all the λ_i's are distinct this in turn implies that

$$a_1 = a_2 = \dots = a_{k-1} = 0,$$

and hence, from (7.6), that $v_k = 0$. But this contradicts the fact that v_k is an eigenvector and hence non-zero by definition. Thus our supposition that $v_1, \dots, v_m$ are dependent is false.

If we apply this theorem to the particular case $m = n$ we see that any matrix with n distinct eigenvalues has n independent eigenvectors and is therefore diagonalizable.

That there exist matrices that are not diagonalizable is easily shown by an example. The simplest example is probably

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

A has only the one independent eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Since a diagonal matrix A' always has the two independent eigenvectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, it follows that A cannot be reduced to A'. To give a complete answer to Question B we would have to consider such matrices as A and see if they can be transformed into some simple form other than the diagonal one. This can be done, leading to the so-called Jordan Canonical Form, but the process is far from simple. The interested student is referred to Halmos, loc. cit.

Problems

7.1 Taking A to be the matrix

$$\begin{pmatrix} 0 & -2 & 2 \\ 2 & -1 & 0 \\ 2 & -2 & 1 \end{pmatrix}$$

of Problem 5.2(c) construct the matrix C for which CAC^{-1} is in diagonal form, and test your value of C by evaluating CAC^{-1}.

7.2 Show by direct calculation that there are no values of a, b, c, d, λ, μ such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

[Hint. cf. Problem 14.4 of Chapter 2.]

7.3 Prove that if either A or B is non-singular then AB and BA have the same characteristic function.
[Hint. $BA = B(AB)B^{-1}$ if $\det B \neq 0$.] (This is also true if both A and B are singular, but a different proof is needed.) It follows that $tr(AB) = tr(BA)$. Prove this directly from the definition of trace in Problem 5.8.

7.4 Let A and B be $n \times n$ matrices each having n distinct eigenvalues.

(a) Prove that A and B commute (i.e., $AB = BA$) if and only if they are simultaneously diagonalizable, that is, there exists a non-singular matrix C such that CAC^{-1} and CBC^{-1} are both diagonal matrices.

(b) Prove that A and B commute if and only if they have the same eigenvectors.

8. Orthogonal Transformations.

In Chapter 2 we saw that the basic properties of a vector space, addition and multiplication by scalars, were preserved under linear transformations. That is, for any linear transformation L we have

$$L(v + w) = Lv + Lw, \quad \text{and} \quad L(cv) = c(Lv).$$

We are now considering vector spaces in which an inner product is defined, and we naturally want to pay special attention to those transformations that preserve inner products, that is, transformations L for which

(8.1) $$(Lu)\cdot(Lv) = u\cdot v.$$

Such transformations are called orthogonal transformations since in particular they preserve orthogonality relations. They also preserve length, that is, $\|Lv\| = \|v\|$; this follows from (8.1) by taking $u = v$.

The theory of orthogonal transformations is concerned almost completely with the case in which the vectors v and the transformed vectors $w = Lv$ are regarded as lying in the same vector space V. In other words, we consider only transformations of V into itself. We are thus entitled to write such expressions as $v\cdot(Lv)$ or to consider the eigenvalue problem $Lv = \lambda v$.

As usual, we restrict out considerations to the finite dimensional case, assuming that an orthonormal basis has been chosen in V, in terms of which vectors can be represented as n x 1 matrices and linear transformations as n x n matrices. A matrix representing an orthogonal transformation is called an orthogonal matrix.

Example 8.1. Problem 9.2 of Chapter 2 consisted in showing that the linear transformation L in V_2 defined by

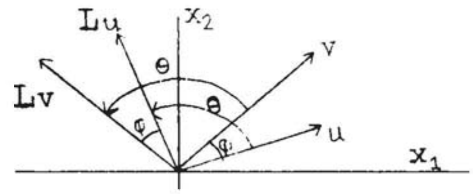


Figure 8.1

$x_1\cos\theta - x_2\sin\theta = y_1$,

$x_1\sin\theta + x_2\cos\theta = y_2$,

is a rotation of the vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ through an angle θ into the vector $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Hence if u and v are two vectors making an angle φ, it follows that Lu and Lv also make an angle φ (Figure 8.1). We also have $\|u\| = \|Lu\|$ and $\|v\| = \|Lv\|$, since the length of a vector is not changed by a rotation. Therefore

$$\begin{aligned}(Lu)\cdot(Lv) &= \|Lu\|\ \|Lv\|\ \cos\varphi \\ &= \|u\|\ \|v\|\ \cos\varphi = u\cdot v.\end{aligned}$$

Thus a rotation in a plane is an example of an orthogonal transformation, and $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is an orthogonal matrix.

The following theorem gives the basic algebraic characterization of an orthogonal matrix.

Theorem 8.1. A matrix R is orthogonal if and only if $R^tR = I$.

Proof. The condition for R to be orthogonal is that

$$(8.2) \qquad (Ru)\cdot(Rv) = u\cdot v$$

for all vectors u and v. (8.2) can be written in matrix form as

$$(8.3) \qquad (Ru)^t(Rv) = u^t v$$

or

$$(8.4) \qquad u^tR^tRv = u^tIv$$

or, finally,

$$u^t(R^tR - I)\, v = 0 \, . \tag{8.5}$$

Obviously if $R^tR - I = 0$ then (8.5) is satisfied for all u and v. Hence if $R^tR = I$ then R is orthogonal, which proves half the theorem.

To prove the other half we introduce the base vectors $e_1,\ldots,e_n$; e_i has its i-th component equal to 1, with all other components zero. For any matrix $A = (a_{ij})$ it is easy to check that $e_i{}^tAe_j = a_{ij}$. Now assume that R is orthogonal. Then (8.5) holds for all choices of u and v; hence in particular for $u = e_i$, $v = e_j$. It follows that the ij-th element of $R^tR - I$ is zero, and since i and j are arbitrary all elements are zero. Hence $R^tR = I$, which proves the other half of the theorem.

Example 8.2. In Example 8.1 we saw that a rotation in V_2 is an orthogonal transformation. Now we shall proceed in the opposite direction and see what we can say in general about an orthogonal transformation in V_2. Such a transformation must have a matrix

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

for which $R^tR = I$; i.e.,

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Hence

$$a^2 + c^2 = 1, \quad b^2 + d^2 = 1, \quad ab + cd = 0 \, .$$

Since $a^2 + c^2 = 1$ we can set

$$a = \cos\theta, \qquad c = \sin\theta,$$

and similarly

$$d = \cos\varphi, \qquad b = \sin\varphi \, .$$

Then $ab + cd = 0$ becomes

$$\sin\varphi \cos\theta + \cos\varphi \sin\theta = 0,$$

or

$$\sin(\varphi + \theta) = 0;$$

so that

$$\varphi = -\theta \text{ or } \varphi = \pi - \theta \, .$$

If $\varphi = -\theta$ then $d = a$, $b = -c$, and

(8.6)
$$R = R_+ = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

If $\varphi = \pi - \theta$ we get

(8.7)
$$R = R_- = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

Thus any 2 x 2 orthogonal matrix has one of these two forms. The transformation $w = R_+v$ is a rotation through an angle θ; the transformation $w = R_-v$ can be obtained by first reflecting in e_1 and then rotating through angle θ. (See Figure 8.2.)

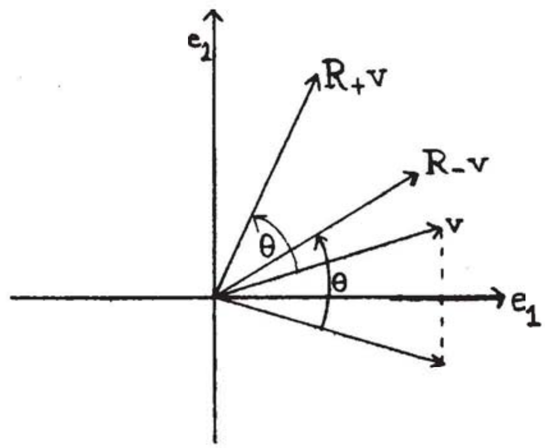


Figure 8.2

Many important properties of orthogonal matrices follow as corollaries to Theorem 8.1.

<u>Corollary 8.1</u>. An orthogonal matrix R is non-singular, and $R^{-1} = R^t$.

<u>Corollary 8.2</u>. The transpose of an orthogonal matrix is orthogonal.

<u>Corollary 8.3</u>. The determinant of an orthogonal matrix is 1 or -1.

<u>Proof</u>. If $R^tR = I$ then

$$(\det R^t)(\det R) = \det I = 1.$$

Since $\det R^t = \det R$, we get

$$(\det R)^2 = 1, \quad \text{or} \quad \det R = \pm 1 .$$

<u>Corollary 8.4</u>. The product of orthogonal matrices is orthogonal.

<u>Proof</u>. If Q, R, S are orthogonal, then

$$
\begin{aligned}
(QRS)^t(QRS) &= S^tR^tQ^tQRS \\
&= S^tR^t(Q^tQ)RS \\
&= S^tR^tIRS \\
&= S^t(R^tR)S \\
&= S^tIS \\
&= S^tS \\
&= I \ .
\end{aligned}
$$

The argument obviously works for the product of any number of factors.

The general structure of an orthogonal matrix is shown by the following theorem.

Theorem 8.2. The column vectors of an orthogonal matrix form an orthonormal set. Conversely any square matrix whose column vectors form an orthonormal set is orthogonal.

Proof. If v_i is the i-th column vector of a matrix R then the ij-th element of R^tR is easily seen to be $v_i \cdot v_j$. The theorem follows readily from this observation.

Corollary 8.5. The above theorem holds also for row vectors.

This follows from Corollary 8.2. It leads to the unexpected conclusion that if the column vectors of a square matrix are an orthonormal set so also are the row vectors.

In Section 7 we also encountered a matrix made up of specified column vectors. There we found that if the columns of C^{-1} were independent eigenvectors of a matrix A, then CAC^{-1} was a diagonal matrix with the eigenvalues of A as its diagonal elements. Now if A is symmetric it has an independent set of n orthogonal eigenvectors (Theorem 5.6 and preceding material) and these can be made orthonormal by dividing each one by its length. It then follows from Theorem 8.2 that the C^{-1} with these column vectors is an orthogonal matrix R, and hence $C = R^{-1} = R^t$. We have therefore proved

Theorem 8.3. If the column vectors of R are an orthonormal set of eigenvectors of a symmetric matrix A, then R is orthogonal and R^tAR is a diagonal matrix with the eigenvalues of A as diagonal elements.

This theorem does not assist us directly in computing eigenvalues, since we need to know the eigenvectors before we can form the matrix R. It serves, however, as a guide in a useful method that enables us to approximate all the eigenvalues of a matrix at the same time. We conclude this chapter with a description of the method and an indication, without precise proof, of why it works.

Consider first a 2 x 2 symmetric matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} .$$

We have seen in Example 8.1 that one type of orthogonal transformation is the rotation

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

If we compute R^tAR we get a matrix

$$B = \begin{pmatrix} a' & b' \\ b' & c' \end{pmatrix}$$

where, in particular

$$b' = b \cos 2\theta - \frac{a-c}{2} \sin 2\theta \ .$$

If A is not already in diagonal form we have $b \neq 0$, and so if we choose θ from

$$\cot 2\theta = \frac{a-c}{2b}$$

we will have $b' = 0$; that is B will be in diagonal form, its diagonal elements will be the eigenvalues, and the rows of R will be the associated eigenvectors.

Next consider the case $n = 3$. For definiteness we will use the specific matrix of Example 6.2, namely

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} .$$

We try the above trick to eliminate the off-diagonal elements, beginning with the largest. That is, we take

$$R_1 = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

In forming R^tAR the upper left 2 x 2 block will be affected just as in the case above. We get θ_1 from $\cot 2\theta_1 = 1/4$, which gives

$\sin\theta_1 = .615$, $\cos\theta_1 = .788$. Straightforward computation then gives

$$A_1 = R_1{}^t A R_1 = \begin{pmatrix} 3.56 & 0 & 1.40 \\ 0 & -.56 & .17 \\ 1.40 & .17 & 0 \end{pmatrix}.$$

Now we go after the 1.40 element, using

$$R_2 = \begin{pmatrix} \cos\theta_2 & 0 & -\sin\theta_2 \\ 0 & 1 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{pmatrix}.$$

Here $\cot 2\theta_2 = (3.56 - 0)/2.80 = 1.271$, $\sin\theta_2 = .327$, $\cos\theta_2 = .945$,

$$A_2 = R_2{}^t A_1 R_2 = \begin{pmatrix} 4.04 & .06 & 0 \\ .06 & -.56 & .16 \\ 0 & .16 & -.48 \end{pmatrix}.$$

At first we seem to be getting nowhere, for introducing new zeros destroys the ones we already had. But on the whole the off-diagonal elements seem to be getting smaller, and this can be shown always to be the case in the following sense: at each step the sum of the squares of the off-diagonal elements is decreased by the squares of the elements that are replaced by zeros. Because of this behavior the process converges to the limit in which all off-diagonal elements are zero, and one can make an estimate as to how many steps will suffice to get within any desired approximation of this limit. If one takes m stages then

$$A_m = R_m{}^t R_{m-1}{}^t \ldots R_2{}^t R_1{}^t A R_1 R_2 \ldots R_{m-1} R_m$$

has (to within this approximation) the desired diagonal structure and the columns of

$$R = R_1 R_2 \cdots R_{m-1} R_m$$

are the eigenvectors.

This method is not well suited for hand computation but it works quite well on a machine. It is known as the Jacobi Method.

Problems

8.1 Show that the following matrices are orthogonal:

$$R = \begin{pmatrix} 2/3 & 2/3 & 1/3 \\ 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \end{pmatrix}, \quad H = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix}.$$

8.2 Show that the matrix R above has the eigenvector (3,1,1) with the associated eigenvalue 1. What does this suggest about the geometric nature of the transformation $w = Rv$? Can you prove your conjecture?

8.3 (a) Prove that transformation (8.6) is a rotation through angle θ. [Hint. Use the polar forms $x = r \cos\beta$, $y = r \sin\beta$.]

(b) Prove that transformation (8.7) is a reflection in the line through the origin making an angle of $\frac{\theta}{2}$ with e_1. [One way of attacking this is to find the eigenvalues and eigenvectors of R_-.]

8.4 Prove that if u is a unit vector then the matrix $A = I - 2uu^t$ is both symmetric and orthogonal.

8.5 Use Theorem 8.3 to prove that if a symmetric matrix A has no negative eigenvalues then there is a symmetric matrix C such that $C^2 = A$. [Hint. If $R^tAR = D$ is a diagonal matrix then D has a square root, i.e., $D = E^2$, where E is also a diagonal matrix. Take $C = RER^t$.]

8.6 Prove that a real eigenvalue of an orthogonal matrix R is either 1 or -1.

[Hint. Apply Theorem 2.2 and the property $\|Rv\| = \|v\|$ to $Rv = \lambda v$.]

For complex eigenvalues see Problem 10.3.

8.7 (a) Prove that if S is both orthogonal and symmetric then $S^2 = I$, and every eigenvalue of S is either 1 or -1.

(b) Show that any such matrix S is of the form R^tJR, where J is a diagonal matrix with ±1's on the diagonal, and R is any orthogonal matrix.

(c) Obtain symmetric orthogonal matrices (not diagonal) by taking the R in (b) to be the R and the H of Problem 8.1.

8.8 Let R be a 3 x 3 orthogonal matrix. The characteristic equation of R is then of odd degree and so has a real root, which by Problem 8.5 is either 1 or -1. Suppose it is 1, and let v_1 be an associated eigenvalue. If u is any vector in the plane orthogonal to v_1 show that Ru is also in this plane. Hence R effects an orthogonal transformation in this plane, and so Example 8.2 applies. Using these facts describe geometrically the transformation $Rv = w$ (there are several possibilities).

Do the same, supposing the real root to have been -1.

9. Quadratic Forms

Several times in this chapter we have had occasion to consider expressions of the type $v^t Av$, where A is a symmetric matrix and v a vector. Such an expression is called a *quadratic form*. If we regard the elements of A as constants and those of v as variables $x_1, x_2, \ldots, x_n$, then $v^t Av$ is a homogeneous second-degree polynomial in the x's - hence the name.* Conversely any such polynomial can be expressed as $v^t Av$, for example

$$2x_1^2 + x_2^2 - x_3^2 + 2x_1x_2 - x_2x_3 = v^t Av,$$

where

$$v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & -1/2 \\ 0 & -1/2 & -1 \end{pmatrix} .$$

Quadratic forms arise in a variety of ways, one of the most important being from consideration of work done on a mechanical system. In Section 4 the relation F = Md was derived for the connection between a force F on a certain type of structure and the resulting displacement d. The work done on the structure by the displacement d is then $W = d \cdot F$, or $W = d^t Md$. We saw that M is a symmetric matrix, so W is thus a quadratic form in d.

*In general, a *form* is simply a homogeneous polynomial, but now-a-days this word is commonly used only in certain special cases.

This simple example is typical of a wide class of physical systems, the so-called "conservative systems in equilibrium" (see Karman and Biot, Mathematical Methods in Engineering, McGraw-Hill Book Co., New York, 1940, Chapter V). The work done on such a system by a small displacement (which may involve changes in several mechanical, electrical, hydraulic, etc. components), represented by an n-component vector v, is given by $W = v^tAv$, where A is an n x n symmetric matrix associated with the system. If the system is in stable equilibrium the value of W will be positive for any displacement d (assuming, of course, that $d \neq 0$). We are thus led to introduce some definitions.

Definition 9.1. A symmetric matrix A and its associated quadratic form v^tAv are said to be

(a) Positive definite if $v^tAv > 0$ for any $v \neq 0$;

(b) Positive semi-definite if $v^tAv \geqq 0$ for any v;

(c) Indefinite if there are v and w such that $v^tAv > 0$, $w^tAw < 0$.

One can also define negative definite and negative semi-definite in analogy with (a) and (b).

An important connection between these concepts and the eigenvalues of the matrix is an immediate consequence of Problem 6.5.

Theorem 9.1. A symmetric matrix is (a) positive definite, (b) positive semi-definite, (c) indefinite, according as

(a) all its eigenvalues are positive,

(b) none of its eigenvalues are negative,

(c) it has both positive and negative eigenvalues.

Proof. Let $v_1,\ldots,v_n$ be an orthonormal basis of eigenvectors of the symmetric matrix A, with associated eigenvalues $\lambda_1,\ldots,\lambda_n$. As in Section 6 we have for any vector v,

$$v = \sum_{i=1}^{n} a_i v_i \quad , \qquad Av = \sum_{i=1}^{n} \lambda_i a_i v_i \ .$$

Then, as in Section 3,

$$v^t Av = v \cdot Av = \sum_{i=1}^{n} \lambda_i a_i^{\,2} \ . \tag{9.1}$$

(This is the statement of Problem 6.5.) Since $a_i^{\,2} > 0$ if $a_i \neq 0$, the statements of the theorem follow at once from this relation. For if no $\lambda_i < 0$, $v^t Av$ can never be negative; if in addition no $\lambda_i = 0$, $v^t Av$ can be zero only if every $a_i = 0$, that is, $v = 0$. And if $\lambda_p > 0$, $\lambda_q < 0$ then $v_p^{\,t} Av_p > 0$, $v_q^{\,t} Av_q < 0$.

One might propose to use this theorem to determine whether or not a given symmetric matrix is positive definite, but in any but the most trivial cases the work involved in determining the eigenvalues would be formidable. Fortunately there is a much simpler criterion, namely: The symmetric nxn matrix $A = (a_{ij})$ is positive definite if and only if the following n inequalities are true,

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \ldots, \det A > 0.$$

This is known as the Routh-Hurwitz criterion. For its proof the reader is referred to G. Hadley, Linear Algebra, Addison-Wesley Publishing Co., Reading, Mass., 1961.

In terms of quadratic forms we can get nice geometric interpretations of the eigenvalues and eigenvectors of a symmetric matrix, of the transformation $Av = w$, and of the computational process of Section 6. Consider the case $n = 2$ and let

$$v = \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Then $v^tAv = ax^2 + 2bxy + cy^2$. We consider the locus defined by $v^tAv = 1$. This is a curve C in the xy-plane, and being of second degree is a conic or some degenerate form of a conic. The point of C in a direction determined by a unit vector u is found by setting $v = ku$ and solving $(ku)^tA(ku) = 1$ for k. We get $k = \dfrac{1}{\sqrt{u^tAu}}$, provided $u^tAu > 0$. If A is positive definite this condition is satisfied for any u, and it follows that C is an ellipse (Figure 9.1a). If A is indefinite we get intersections for some values of u but not for others, and in this case C is a hyperbola (Figure 9.1b).

To get an interpretation of the linear transformation $Av = w$ we must consider the tangents to C. If we think of C

$$2x^2 + 4xy + 5y^2 = 1.$$

$$A = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}.$$

$$\lambda_i = 1, 6.$$

$$v_i = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{30}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

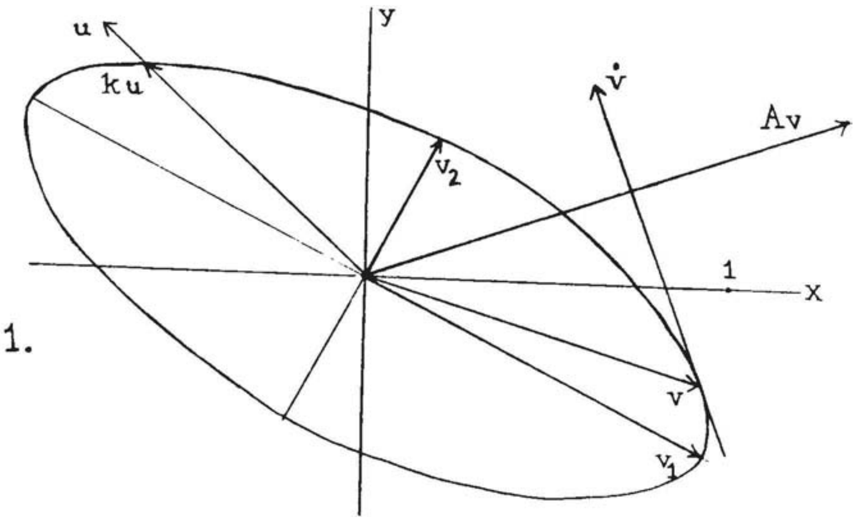


Figure 9.1 a

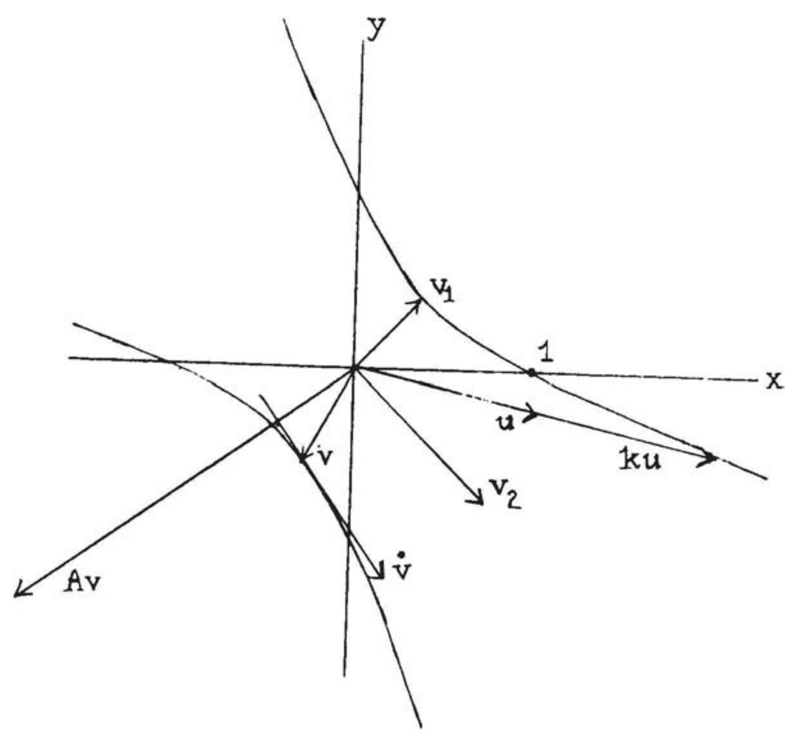
Figure 9.1 b

$$x^2 + 6xy + y^2 = 1.$$

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}.$$

$$\lambda_i = 4.-2.$$

$$v_i = \frac{1}{2\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$



as the path of a moving particle whose coordinates $(x(t), y(t))$ are functions of time then (cf. Thomas, Section 12-5) the velocity vector, which is tangent to C, is $\dot{v} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}$, where the dot indicates differentiation with respect to time. Now differentiate with respect to time the equation

$$ax^2 + 2bxy + cy^2 = 1$$

of C. Dividing by 2 gives

$$a\dot{x}x + b\dot{x}y + b\dot{y}x + c\dot{y}y = 0,$$

or

$$\dot{x}(ax + by) + \dot{y}(bx + cy) = 0,$$

or finally, in matrix form,

$$\dot{v}^t Av = 0$$

Geometrically, this says that Av is perpendicular to the line tangent to C at the tip of v (see Figure 9.1). In particular, if v is an eigenvector of A then v and Av have the same direction, and so <u>an eigenvector of A which intersects C is normal to</u> C. The eigenvectors are therefore identified with the axes of the conic C. How about the eigenvalues? If $Av = \lambda v$ and $v^t Av = 1$, then $\lambda v^t v = 1$, or $\|v\| = \lambda^{-1/2}$. Hence <u>the half-length of the axis is the rec iprocal square root of the associated eigenvalue</u>.

It is now easy to see what is happening in the repeated multiplication process of Section 6. Starting with a vector

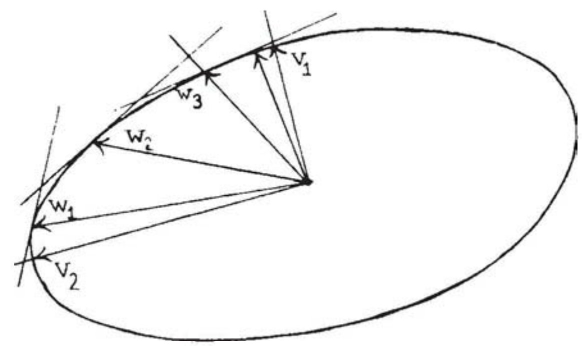


Figure 9.2

w_1, and letting $w_2 = Aw_1$, $w_3 = Aw_2, \ldots$, except for scalar multipliers, we get the situation shown in Figure 9.2. The successive vectors approach the minor axis, that is, the axis associated with the maximum eigenvalue.

These geometric considerations can be extended to values of n greater than 2. For n = 3, $v^tAv = 1$ is the equation of a quadric surface Q, which is an ellipsoid if A is positive definite. The vector Av is perpendicular to the plane tangent to Q at the tip of v. The eigenvectors of A are the axes of Q, and the half-length of an axis is again $\lambda^{-1/2}$. For still larger values of n we can only proceed by analogy but the geometric language is still illuminating at times. For indefinite matrices the picture gives no information about those vectors that do not intersect Q - in particular an eigenvector associated with a negative eigenvalue has no simple distinguishing property. All-in-all one must accept these geometric interpretations as no more than what they are, aids to the understanding and possible prediction of algebraic properties. In the final analysis the proofs must be carried out by algebra.

A standard problem in analytic geometry is that of changing axes so that the equation of a given conic assumes a simple form. In particular, for a conic with center at the origin this involves a rotation of axes to make the new coordinate axes coincide with the axes of the conic (cf. Thomas, Section 9-9). Let us consider this problem for a general quadratic form $v^t A v$.

We adopt the point of view of Chapter 2, Section 15, namely that v is a column vector whose elements are the components of some physical vector $\vec{v}$ with respect to a specific basis in the space V of $\vec{v}$. (For n = 3 think of $\vec{v}$ as an arrow from the origin and v as the column vector

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

giving the components of $\vec{v}$ in a particular coordinate system.) A change of basis in V leaves $\vec{v}$ unchanged but changes v to $v' = Cv$, where C is a non-singular matrix. We shall find it convenient to set $S = C^{-1}$ and write the relation between v and v' in the form

(9.2) $$v = Sv' .$$

Now the equation $v^t A v = 1$ can be though of as defining a locus Q, in the vector space V, that is independent of the choice of basis. Hence when we change the basis by (9.2) the new equation becomes $v'^t S^t A S v' = 1$, or $v'^t A' v' = 1$, where

(9.3) $$A' = S^t A S .$$

Thus (9.3) tells us how the <u>matrix of a quadratic form</u> is affected by a change of basis (9.2). In Chapter 2 we found that the effect of (9.2) on the <u>matrix of a transformation</u> is given by

(9.4) $$A' = S^{-1}AS \ .$$

It is evident that the matrix behaves differently in these two aspects unless we restrict ourselves to changes of bases such that $S^t = S^{-1}$. Since this equation is equivalent to $S^tS = I$ we see that our restriction is a very natural one, namely we consider only orthogonal changes of bases.

We can now apply Theorem 8.3 to reduce our quadric Q to a standard form.

<u>Theorem 9.2</u>. If an orthonormal set of eigenvectors of the symmetric matrix A is taken as a new basis, the quadratic form v^tAv reduces to $\sum_{i=1}^{n} \lambda_i x_i{}^2$.

<u>Proof</u>. Theorems 8.3 and 7.2 tell us that with the new basis the matrix A is replaced by

$$A' = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} .$$

We see at once that $v^tA'v = \sum_{i=1}^{n} \lambda_i x_i{}^2$.

We conclude this section with a useful application of some of these concepts.

In many cases one meets the eigenvalue problem in the more general form

$$Av = \lambda Bv , \tag{9.5}$$

where A and B are given n x n matrices. If B is non-singular, as is usually the case, (9.5) can be replaced by

$$B^{-1}Av = \lambda v \tag{9.6}$$

and the usual eigenvalue methods applied to the matrix $B^{-1}A$. (But see Problem 9.7 for an approach involving less computation.) However, even if A and B are both symmetric $B^{-1}A$ need not be, and so we cannot conclude that the solutions of (9.5) have the desirable properties we expect when dealing with symmetric matrices. That trouble can actually occur is easily seen from an example, for if $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ then

$B^{-1}A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has $\pm i$ as eigenvalues. The following theorem tells us that in one commonly occurring case this trouble does not arise.

Theorem 9.3. If A and B are symmetric and if B is positive definite then the eigenvalues λ of (9.5) are real and there is a basis of eigenvectors.

Proof. Let R be an orthogonal matrix such that $R^tBR = D$ is a diagonal matrix. The diagonal elements of D are the eigenvalues of B and hence (Theorem 9.1) they are all positive.

Let E be the diagonal matrix having as its i-th diagonal element the reciprocal square root of the corresponding element of D. Then $R^tBR = E^{-2}$ and

(9.7) $$ER^tBRE = I \quad .$$

We now let

(9.8) $$v = REw, \quad \text{or} \quad w = E^{-1}R^{-1}v \ .$$

Substituting for v in (9.5) and multiplying on the left by ER^t gives, in virtue of (9.7)

(9.9) $$ER^tAREw = \lambda w.$$

Now $(ER^tARE)^t = E^tR^tA^tRE^t = ER^tARE$ since A and E are symmetric, and so λ and w are eigenvalues and eigenvectors of a symmetric matrix and hence the λ's are real and there is a basis of the w's. But (9.8) is a coordinate transformation, since RE is non-singular, and so the v's also form a basis (but not necessarily an orthogonal one).

Problems

9.1 If B is any n x n matrix and $A = 1/2(B + B^t)$ show that A is symmetric and that $v^tBv = v^tAv$. Hence nothing is gained by considering quadratic forms defined by non-symmetric matrices.

9.2 What can you say about -A if A is

(a) positive definite,

(b) negative semi-definite,

(c) indefinite.

9.3 Prove the following:

(a) For any matrix A (not necessarily square), A^tA is positive semi-definite.

(b) If A is non-singular, then A^tA is positive definite.

9.4 If A is an indefinite 2 x 2 matrix what is the geometric significance of those vectors u for which $u^tAu = 0$?

9.5 Describe the quadric surface $v^tAv = 1$ if the 3 x 3 matrix A has

(a) two positive and one negative eigenvalues,

(b) one positive and two negative eigenvalues.

[See Thomas, Section 13-10.]

9.6 For n = 3 what can you say about the quadric surface if the matrix has two equal eigenvalues? What does this imply about the associated eigenvectors, and how is this related to Corollary 5.2?

9.7 Prove that if A is symmetric, then $\frac{d}{dt}(v^tAv) = 2\dot{v}^tAv$.

[Hint. $v^tAv = \sum_{i,j=1}^{n} x_i a_{ij} x_j$; and if A is symmetric

$x_i a_{ij} \dot{x}_j = \dot{x}_j a_{ji} x_i$.]

9.8 Show that the generalized eigenvalue equation $Av = \lambda Bv$ can be solved in the same manner as the standard equation, i.e., by solving for λ the equation

$$\det(\lambda B - A) = 0$$

and for each root λ finding an independent basis for the solutions of

$$(\lambda B - A)v = 0.$$

9.9 By the method of Problem 9.8 solve the generalized eigenvalue problem in each of the following cases:

(a) $A = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

(b) $A = \begin{pmatrix} 2 & -1 & -1 \\ 4 & 1 & -1 \\ 4 & 4 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

9.10 Derive conditions on a, b, c that will insure that $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is positive definite.

9.11 Given a positive definite symmetric matrix B, for any two vectors u and v we define $u \circ v = u^t Bv$. Show that $u \circ v$ satisfies the five properties of Definition 2.1 and is therefore an inner product.

9.12 Let A and B be symmetric and B positive definite. If $C = B^{-1}A$ show that for any vectors u and v, $u \circ Cv = Cu \circ$ where $\circ$ is defined as in Problem 9.11.

9.13 Using the results of Problems 9.11 and 9.12, modify the proofs of Theorems 5.4 and 5.5 to prove that under the conditions of Theorem 9.3:

(a) The eigenvalues of $Av = \lambda Bv$ are real;

(b) If v_1 and v_2 are eigenvectors of $Av = \lambda Bv$ associated with different eigenvalues, then $v_1 \circ v_2 = 0$.

10. Vector Spaces Over the Complex Field

In Section 5 we found it necessary to introduce complex numbers into our considerations. We did our best to avoid this complication by concentrating our attention on symmetric matrices, but obviously we cannot expect that we shall encounter only these special cases. There is therefore some value in systematically investigating the complex case.

To start at the beginning, we go back to the definition of a vector space, Definition 3.1 of Chapter 2. This definition involves certain operations with real numbers, which we call scalars. Now instead of real numbers let our scalars be complex numbers. We have seen (Section 2 of Chapter 4) that the complex numbers form an algebraic field, and from this it follows that all the developments of Chapter 2 apply to a vector space with complex scalars.

We first run into trouble, as we saw in Section 5, in connection with the inner product. We cannot afford to allow $v \cdot v$ to be zero if $v \neq 0$, as is the case if $v = (1,i)$, since this would ruin many of our standard processes, such as the Schmidt orthogonalization, that depend on division by $v \cdot v$. The trick used in Section 5 suggests that we modify the definition of the inner product of two n-tuples by defining

$$(10.1) \quad (a_1,\ldots,a_n)\cdot(b_1,\ldots,b_n) = \sum_{i=1}^{n} \bar{a}_i b_i \ ,$$

where $\bar{a}_i$ denotes, as usual, the complex conjugate of a_i. Then

$$(a_1,\dots,a_n)\cdot(a_1,\dots,a_n) = \sum_{i=1}^{n}\bar{a}_i a_i = \sum_{i=1}^{n}|a_i|^2 \geqq 0,$$

where equality holds only if each

$a_i = 0$, that is, $(a_1,\dots,a_n) = 0$.

This trick manages to satisfy property 5 of the inner product (See Definition 2.1) but it forces modifications of properties 2 and 3. These become

2. $u\cdot v = \overline{v\cdot u}$;

3. For any scalar c, $c(u\cdot v) = u\cdot(cv) = (\bar{c}u)\cdot v$.

Thus the symmetry of the original definition is unfortunately lost, but there seems to be no way to avoid this.

The proofs of all but one of the formulas in Section 2 and 3 can be obtained in essentially the same way as before. We must, however, take account of the non-commutativity of the inner product. Thus, for example, (3.5) is not correct as it stands but must be changed to

$$a_i = \frac{v_i\cdot v}{v_i\cdot v_i} \quad .$$

The one theorem that requires special treatment is the Schwarz Inequality, Theorem 2.3. This needs a new statement, namely

$$|v\cdot w|^2 \leqq (v\cdot v)(w\cdot w) \ ,$$

and a somewhat more elaborate proof. Instead of (2.3) we start with

(10.2) $$0 \leqq (tv + cw)\cdot(tv + cw),$$

where t is a real variable and c a scalar to be selected later. Expanding (10.2) gives

(10.3) $$0 \leqq (v\cdot v)t^2 + [c(v\cdot w) + \bar{c}(w\cdot v)]t + |c|^2(w\cdot w).$$

We wish to choose c so that the expression in the brackets becomes $2|v\cdot w|^2$. This will occur if $c = (\overline{v\cdot w})$, for then

$$c(v\cdot w) = (\overline{v\cdot w})(v\cdot w) = |v\cdot w|^2$$

and

$$\bar{c}(w\cdot v) = (v\cdot w)(\overline{v\cdot w}) = |v\cdot w|^2 .$$

(10.3) then becomes

$$0 \leqq (v\cdot v)t^2 + 2|v\cdot w|^2 t + |v\cdot w|^2(w\cdot w).$$

With this inequality taking the place of (2.3) the rest of the proof follows as before.

The inner product of two column vectors, u and v, now becomes $\bar{u}^t v$ instead of $u^t v$, where $\bar{u}$ is the n x 1 matrix whose elements are the conjugates of those of u. It turns out that in dealing with complex vector spaces the conjugate transpose $\bar{A}^t$ of a matrix A is generally of more importance than A^t. We therefore give $\bar{A}^t$ an abbreviation, A^*, and a name, the adjoint of A. (Refer to the comment on page 2.145 of Volume 1.)

We can now go systematically through the theory of Sections 5 to 9, making the appropriate changes indicated above, and essentially duplicating the results and their proofs.

Instead of symmetric matrices, for which $A^t = A$, we consider <u>Hermitian</u> matrices, defined by $H^* = H$; instead of orthogonal matrices, $R^tR = I$, we have <u>unitary</u> matrices, $U^*U = I$. As an important example, Theorem 8.3 becomes: If the column vectors of U are an orthonormal basis of eigenvectors of a Hermitian matrix H then U^*HU is a diagonal matrix with the eigenvalues of H as diagonal elements.

Problems

10.1 Which of the following matrices are (i) Hermitian, (ii) unitary?

$$\text{(a)} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \text{(c)} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix},$$

$$\text{(d)} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \text{(e)} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}, \quad \text{(f)} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

10.2 State and prove the analogs of Theorems 5.4 and 5.5 for the complex case.

10.3 A <u>Hermitian</u> <u>form</u> is defined to be v^*Hv, where H is a Hermitian matrix. Show that v^*Hv is a real number for any v, and hence that Definition 9.1 can be applied to Hermitian forms. Prove the analog of Theorem 9.1.

10.4 (a) Prove that every eigenvalue of a unitary matrix has absolute value 1. [cf Problem 8.5] Since an orthogonal matrix is merely a unitary matrix whose elements are real, the same property holds for any orthogonal matrix.

(b) Prove the analog of Theorem 5.5 for unitary matrices. [Hint. First use the result of (a) to show that if $Uw = \mu w$ then $U^* w = \bar{\mu} w$.]

10.5 Following (2.2) and (10.1) we define

$$f \cdot g = \int_a^b \overline{f(x)} g(x) dx ,$$

where f and g are complex valued functions of a real variable x, continuous on the interval $a \leqq x \leqq b$.

(a) Show that on the interval $0 \leqq x \leqq 2\pi$ the functions $\varphi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$, $n = 0, \pm 1, \pm 2, \cdots,$ are orthonormal, i.e.,

$$\varphi_n \cdot \varphi_n = 1, \qquad \varphi_n \cdot \varphi_m = 0 \quad \text{if } n \neq m.$$

(b) Given the function $f(x) = x$, show that

$$\varphi_n \cdot f = \begin{cases} \pi \sqrt{2\pi} \text{ if } n = 0 \\ -\frac{1}{in} \sqrt{2\pi} \text{ if } n \neq 0. \end{cases}$$

[Use $\int x e^{ax} dx = (1/a) x e^{ax} - (1/a^2) e^{ax}$.]

(c) Using the method of Example 3.4 show that for any $N \geqq 1$, the function $f(x) = x$ is approximated on the interval $0 \leqq x \leqq 2\pi$ by the function

$$F_N(x) = \pi - \sum_{n=1}^{N} \frac{e^{inx} - e^{-inx}}{ni}$$

$$= \pi - 2 \sum_{n=1}^{N} \frac{\sin nx}{n} .$$

(d) Show that

$$\|x - F_N(x)\|^2 = (x - F_N(x))\cdot(x - F_N(x))$$

$$= 4\pi\left(\frac{\pi^2}{6} - \sum_{n=1}^{N}\frac{1}{n^2}\right) .$$

This example illustrates an important approach to the theory of Fourier Series.

11. Self-Adjoint Operators

In Section 7 we considered the Eigenvalue Problem $Lv = \lambda v$ in a general vector space V, L being any linear operator from V to V. In the finite-dimensional case L can be represented by a matrix A, and we are then in a position to define such concepts as Hermitian and unitary matrices and to derive their many properties. The question naturally arises as to whether these concepts can be applied to general linear operators L without the intervention of matrices; if so, we could, in particular, apply them to infinite-dimensional function spaces. We shall now see that this can indeed be done. We continue to use complex numbers as scalars, as in the preceding section.

To serve as a guide we first find a way to characterize the adjoint of a matrix A without referring to its rows and columns. If u and v are any vectors, then u^*Av is a scalar, and hence its adjoint is just its conjugate, thus

$$(u^*Av)^* = \overline{u^*Av} = \overline{u\cdot(Av)} = (Av)\cdot u \quad .$$

But also

$$(u^*Av)^* = v^*A^*u = v\cdot(A^*u).$$

Hence

$$v\cdot(A^*u) = (Av)\cdot u \ .$$

This leads us to make the following definitions: Two linear operators L_1 and L_2 are said to be adjoint to each other if $v\cdot(L_1u) = (L_2v)\cdot u$ for every pair of vectors u, v. An operator L is self-adjoint if $v\cdot(Lu) = (Lv)\cdot u$ for every pair of vectors u,v. Plainly, "self-adjoint" is the abstract equivalent of "Hermitian."

For questions of existence and uniqueness of the adjoint of a given operator we refer the reader to Halmos, loc. cit. Our only concern here is to indicate the way in which many of the properties of Hermitian matrices extend to self-adjoint operators.

The properties that do not carry over directly are those that depend on the finite-dimensionality of the vector space, e.g., the characteristic equation, the computational methods of solving the eigenvalue problem, and the existence of a basis of eigenvectors. But the theorems on the reality of the eigenvalues and the orthogonality of eigenvectors (Theorems 5.4 and 5.5) can be proved just as before. Quadratic forms $v\cdot(Lv)$ can be introduced, and classified as in Definition 9.1. Only the last part of the proof of Theorem 9.1 makes no use of a basis, so all we can salvage from this theorem is that a positive definite operator has no negative or zero eigenvalues.

A unitary operator U is a linear operator $Uv = w$ from

V to V such that $(Uv)\cdot(Uw) = v\cdot w$ for all vectors v and w (cf Section 8). If U and U_1 are adjoint operators then

$$(Uv)\cdot(Uw) = [U_1(Uv)]\cdot w = [(U_1U)v]\cdot w,$$

and so if $U_1U = I$ then U is unitary. It is left to the student (Problem 11.1) to prove the converse, namely, that if U is unitary then $U_1U = I$, and thereby to get the generalization of Theorem 8.1.

Example 11.1. Let V be the space C^∞ of functions of x defined on an interval $a \leqq x \leqq b$ and having derivatives of arbitrarily high order. To simplify the discussion we consider only real functions; then the inner product is defined as in (2.2)

$$f\cdot g = \int_a^b f(x)g(x)dx .$$

If f is a function in C^∞ so is D^2f, defined by $D^2f = \frac{d^2}{dx^2} f(x)$. Thus D^2 is an operator from C^∞ to C^∞, and it is well known to be linear. Let us see if D^2 is self-adjoint. We have

$$f\cdot D^2g = \int_a^b f(x)g''(x)dx .$$

We integrate this by parts, taking $f(x) = u$, $g''(x)dx = dv$, to get

$$f\cdot D^2g = [f(x)g'(x)]_a^b - \int_a^b f'(x)g'(x)dx ;$$

another integration by parts gives

$$f\cdot D^2 g = [f(x)g'(x) - f'(x)g(x)]_a^b + \int_a^b f''(x)g(x)dx$$

$$= [f(x)g'(x) - f'(x)g(x)]_a^b + (D^2 f)\cdot g \ .$$

If it were not for the expression in brackets we could say that D^2 is self-adjoint. To get rid of the superfluous material we put a restriction on the functions, admitting only those that assume the value 0 at $x = a$ and $x = b$. In this restricted space D^2 is self-adjoint.

The quadratic form $f\cdot D^2 f$ is

$$f\cdot D^2 f = \int_a^b f(x)f''(x)dx = [f(x)f'(x)]_a^b - \int_a^b f'(x)f'(x)dx$$

$$= -\int_a^b (f'(x))^2 dx \leqq 0.$$

As in Example 2.2 we can show that the equality sign can hold only if $f'(x)$ is identically zero on the interval $a \leqq x \leqq b$. Hence $f(x)$ is a constant, and since $f(a) = 0$, $f(x)$ is identically zero. Thus $f\cdot D^2 f$ is never positive and is zero only when $f = 0$; that is, D^2 is negative definite.

Since D^2 is self-adjoint its eigenvalues are real; since it is negative definite its eigenvalues are negative. Let us compute them. We have to solve

(11.1) $$D^2 f = \lambda f \quad \text{or} \quad \frac{d^2 f}{dx^2} - \lambda f = 0$$

subject to the two point boundary condition (cf. page 1.42)

$$f(a) = f(b) = 0.$$

it simplifies our results, without changing them in any essential manner, if we assume $a = 0$, $b = \pi$.

If $\lambda > 0$ the general solution of (11.1), as we shall see in the next chapter, is

$$f(x) = c_1 e^{\sqrt{\lambda}x} + c_2 e^{-\sqrt{\lambda}x} .$$

The boundary conditions give

$$0 = c_1 + c_2,$$

$$0 = c_1 e^{\pi\sqrt{\lambda}} + c_2 e^{-\pi\sqrt{\lambda}} .$$

These obviously have only the trivial solution $c_1 = 0$, $c_2 = 0$, so λ cannot be an eigenvalue.

If $\lambda = 0$ the general solution is $c_1 + c_2 x$, which again gives $c_1 = c_2 = 0$ from the boundary conditions.

Hence we must have $\lambda < 0$, as we have already seen. If we set $\lambda = -k^2$ the solution of the differential equation is

$$f(x) = c_1 \sin kx + c_2 \cos kx,$$

and we get

$$0 = \qquad c_2 ,$$

$$0 = c_1 \sin k\pi + c_2 \cos k\pi.$$

We get a non-trivial solution if k is chosen so that $\sin k\pi = 0$, i.e., if k is an integer n. Hence the eigenvalues of D^2 are $\lambda = -n^2$, $n = 1,2,\ldots$. With the eigenvalue $-n^2$ is associated only one independent eigenvector $\sin nx$.

These eigenvectors are orthogonal, for

$$\int_0^{\pi} \sin nx \sin mx \, dx = 0 \quad \text{if} \quad n \neq m.$$

(cf. Problem 6.10 of Chapter 4).

This example illustrates the beginning of the theory of self-adjoint differential equations, a topic of great importance in many branches of applied mathematics.

Problems

11.1 (a) By definition, an operator L is zero if $Lv = 0$ for every v in V. Prove that if $(Lv)\cdot w = 0$ for every v and w in V, then $L = 0$. [Hint. Since w is arbitrary, take $w = Lv$.]

(b) Prove that if M is a linear operator such that $(Mv)\cdot w = v\cdot w$ for every v and w in V, then $M = I$. [Hint. Take $M - I$ for the L in (a).]

(c) Prove that if U is unitary and U_1 is adjoint to U then $U_1U = I$.

11.2 Let V consist of all infinitely differentiable functions $f(x)$ on $0 \leqq x \leqq 2\pi$ satisfying the boundary conditions $f(0) = f(2\pi)$, $f'(0) = f'(2\pi)$, $f''(0) = f''(2\pi), \ldots$

$$f^{(n)}(0) = f^{(n)}(2\pi), \ldots \quad .$$

Derive the following properties of the operator D^2.

(a) D^2 is self-adjoint and negative semi-definite.

(b) D^2 has the eigenvalue 0 with one independent eigenvector, 1.

(c) The other eigenvalues of D^2 are $-n^2$, $n = 1,2,\ldots$

(d) The eigenvalue $-n^2$ has two independent associated eigenvectors, which can be taken to be either e^{inx}, e^{-inx} or $\sin nx$, $\cos nx$.

(e) By virtue of the general theory, eigenvectors associated with different eigenvalues are orthogonal. Show that e^{inx} and e^{-inx} are orthogonal, and also $\sin nx$ and $\cos nx$.

(f) How does this problem relate to Problem 10.5?

11.3 Show that the linear operator

$$L = a(x)D^2 + b(x)D + c(x)I,$$

where a,b,c are functions in the vector space V of Problem 11.2, is self-adjoint, neglecting boundary conditions, if $b(x) = a'(x)$. If this relation holds the needed boundary conditions are of the same type as in Example 11.1.

11.4 Prove that if U is length-preserving it is unitary. That is, if $\|Uv\| = \|v\|$ for every vector v, then $(Uv)\cdot(Uw) = v\cdot w$ for every v and w. [Hint. Consider $\|v + w\|$.]

11.5 In Example 11.1 we can use the method of Chapter 2, Problem 1.7, to approximate the differential equation

$D^2 f = \lambda f$ by a system of linear equations. Show that this leads to a matrix eigenvalue problem $Av = h^2 \lambda v$, and determine the form of the matrix A. Can you prove that A is negative definite?

CHAPTER 6

Linear Differential Equations With Constant Coefficients

1. Two First Order Homogeneous Simultaneous Differential Equations.

A large number of physical processes give rise, on translation into mathematical terms, to a system of two first order differential equations of the form

$$(1.1)\qquad \begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 , \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 , \end{aligned}$$

together with initial conditions

$$(1.2)\qquad x_1(t_0) = X_1 , \; x_2(t_0) = X_2 .$$

We call (1.1) a system of homogeneous linear equations. They are a special case of equations (3.1) of Chapter 1 where the functions $f_i(t,x_1,\ldots,x_n)$ are linear in the dependent variables x_1 and x_2, with constant coefficients $a_{i,j}$, and do not involve the independent variable t.

Example 1.1. The Simple Harmonic Oscillator.

A simple physical system which possesses the properties of "inertia" and "stiffness" gives rise to a system of equations of the form (1.1). Figure 1.1 shows an inertia element

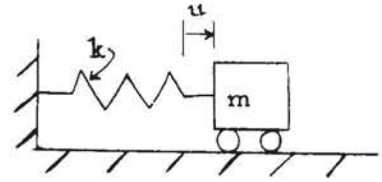


Figure 1.1

of mass m connected to a stiffness element which has a spring constant k. We assume an ideal physical model in which the mass m is perfectly rigid, the spring has negligible mass (compared to m), and there is no friction. Initially, the system is excited by giving the mass m a displacement $u(0) = U_0$ and velocity $v(0) = V_0$. Figure 1.1 shows such a system at any time t with a displacement u from the equilibrium position. If the spring is assumed to be linear, then a change in the length u gives rise to a force ku whose direction is opposite to that of the displacement. Hence, Newton's law and the velocity-displacement relation give the equations

$$(1.3) \qquad \begin{cases} m\,\dfrac{dv}{dt} = -\,ku\,, \\[2ex] \dfrac{du}{dt} = v\,, \end{cases} \qquad \text{or} \qquad \begin{cases} \dfrac{dv}{dt} = -\,\dfrac{k}{m}\cdot u\,, \\[2ex] \dfrac{du}{dt} = v\,, \end{cases}$$

with initial conditions

$$(1.4) \qquad u(0) = U_0\,,\ v(0) = V_0.$$

Equations (1.3) are seen to be a special case of (1.1) with $a_{11} = 0$, $a_{12} = -\frac{k}{m}$, $a_{21} = 1$, and $a_{22} = 0$.

The electrical circuit shown in Figure 1.2 also gives rise to a set of equations of the form (1.3). Two elements having inductance L and capacitance C are connected in series. Initially the capacitor has a charge $Q(0) = Q_0$

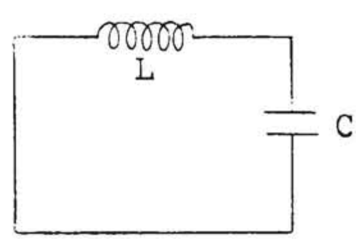


Figure 1.2

and a current $I(0) = I_0$ flows in the circuit. Again we assume an ideal physical model in which the inductor L has no resistance or capacitance and the capacitor C possesses no resistance or inductance. The "inertia" of the system arises from the inductance L, for a change of current is opposed by a voltage $L \frac{dI}{dt}$. The capacitor C is the "stiffness" element where, if the charge Q is assumed proportional to the voltage V across the capacitor, $Q = CV$. Kirchhoff's law for the circuit together with the current-charge relation give the equations

$$(1.5) \qquad \begin{cases} L \frac{dI}{dt} + \frac{1}{C} Q = 0 , \\ \frac{dQ}{dt} = I , \end{cases} \text{or} \begin{cases} \frac{dI}{dt} = - \frac{1}{LC} Q , \\ \frac{dQ}{dt} = I , \end{cases}$$

with initial conditions

$$(1.6) \qquad Q(0) = Q_0 , \quad I(0) = I_0 \quad .$$

Again equations (1.5) are a special case of (1.1) with $a_{11} = 0$, $a_{12} = - \frac{1}{LC}$, $a_{21} = 1$, and $a_{22} = 0$. Comparison of equations (1.5) and (1.6) with (1.3) and (1.4) makes it clear that the electrical and mechanical systems are analogous if we identify corresponding elements as $L \to m$, $\frac{1}{C} \to k$, $I \to v$, and $Q \to u$. The mathematical description is the same in both cases and is the basis of electro-mechanical analogies.

If x_1 and x_2 are elements of a column vector x then $x = \begin{matrix} x_1 \\ x_2 \end{matrix}$, and if the elements are functions of time t then by

$\frac{dx}{dt}$ we mean

(1.7) $$\frac{dx}{dt} = \begin{pmatrix} \frac{dx_1}{dt} \\ \\ \frac{dx_2}{dt} \end{pmatrix} .$$

Thus, (1.1) can be written in vector form as

(1.8) $$\frac{dx}{dt} = Ax ,$$

where A is the matrix of coefficients

(1.9) $$A = \begin{pmatrix} a_{11} & a_{12} \\ \\ a_{21} & a_{22} \end{pmatrix} .$$

If x were a single dependent variable and A a single constant, then from Chapter 1 the solution x(t) of (1.8) would be of exponential form. Hence, one might expect the solution of (1.8) to be of exponential form even when the symbols have the above vector meaning. For example, to solve (1.5) we assume that I and Q have the form

(1.10) $$I = I_n e^{\lambda t} , \quad Q = Q_n e^{\lambda t} ,$$

where I_n and Q_n are constants to be determined from the initial conditions and λ is a constant which must be evaluated. When equations (1.10) are substituted in (1.5) we obtain

$$\lambda I_n e^{\lambda t} = -\frac{1}{LC} Q_n e^{\lambda t} ,$$

(1.11)

$$\lambda Q_n e^{\lambda t} = I_n e^{\lambda t} ,$$

which if $e^{\lambda t}$ is removed can be written as

$$0I_n - \frac{1}{LC} Q_n = \lambda I_n,$$

(1.12)

$$I_n + 0Q_n = \lambda Q_n .$$

The matrix form of (1.12) is

(1.13) $$Ax = \lambda x ,$$

where the matrix A and the column vector x are

$$A = \begin{pmatrix} 0 & -\frac{1}{LC} \\ 1 & 0 \end{pmatrix} , \quad x = \begin{pmatrix} I_n \\ Q_n \end{pmatrix} .$$

It is apparent from (1.13) that <u>the unknown parameter</u> λ <u>is an eigenvalue of the matrix A</u> with I_n and Q_n the components of the corresponding eigenvector. From the previous work of Chapter 5 we can solve for the eigenvalues λ by writing (1.13) in the form

(1.14) $$(\lambda I - A)x = 0 .$$

This system of homogeneous equations, and hence (1.12), have a non-trivial solution if and only if $\det(\lambda I - A) = 0$, so that

(1.15) $$\begin{vmatrix} \lambda & \frac{1}{LC} \\ -1 & \lambda \end{vmatrix} = \lambda^2 + \frac{1}{LC} = 0 .$$

Equation (1.15) is <u>the characteristic equation</u> of the matrix A, with two roots

(1.16) $$\lambda_1 = +i\omega$$

and

(1.17) $$\lambda_2 = -i\omega \ ,$$

where

(1.18) $$\omega = \sqrt{\frac{1}{LC}} \ .$$

Substitution of the eigenvalue λ_1 into either equation of (1.12) gives a relation between the corresponding constants I_1, Q_1 so that the eigenvector $x^{(1)}$ is

(1.19) $$x^{(1)} = \begin{pmatrix} i\omega Q_1 \\ \\ Q_1 \end{pmatrix} .$$

In a similar way, (1.12) gives a relation between the constants I_2, Q_2 corresponding to the eigenvalue λ_2 so that the eigenvector $x^{(2)}$ is

(1.20) $$x^{(2)} = \begin{pmatrix} -i\omega Q_2 \\ \\ Q_2 \end{pmatrix} .$$

Thus, there are two solutions of the form (1.10), for each eigenvalue gives rise to a solution satisfying (1.11). With $\lambda_1 = i\omega$ we obtain

(1.21) $$I = i\omega Q_1 e^{i\omega t}\ ,\ Q = Q_1 e^{i\omega t},$$

and with $\lambda_2 = -i\omega$ we get

(1.22) $$I = -i\omega Q_2 e^{-i\omega t},\ Q = Q_2 e^{-i\omega t}\ .$$

If we add solutions (1.21) and (1.22) we obtain

(1.23) $$I = i\omega Q_1 e^{i\omega t} - i\omega Q_2 e^{-i\omega t}\ ,$$ $$Q = Q_1 e^{i\omega t} + Q_2 e^{-i\omega t}\ ,$$

and substitution of expressions (1.23) into the original gov- erning equations (1.5) gives

(1.24) $$\frac{d}{dt}\left\{ i\omega Q_1 e^{i\omega t} - i\omega Q_2 e^{-i\omega t}\right\} \overset{?}{=} -\omega^2 \left\{ Q_1 e^{i\omega t} + Q_2 e^{-i\omega t}\right\},$$ $$\frac{d}{dt}\left\{ Q_1 e^{i\omega t} + Q_2 e^{-i\omega t}\right\} \overset{?}{=} \left\{ i\omega Q_1 e^{i\omega t} - i\omega Q_2 e^{-i\omega t}\right\}.$$

When the differentiations on the left hand side of equations (1.24) are carried out, we observe that the two sides are indeed the same, the question marks can be removed, and the governing differential equations are satisfied. We conclude that the sum (1.23) of the two solutions (1.21) and (1.22) is itself a solution. This important result shows that by adding the solutions corresponding to each eigenvector we construct a new solution (1.23). The constants Q_1 and Q_2 are obtained from the initial conditions (1.6); thus, if initially the capacitor has charge Q_0 and no current is flowing, $I_0 = 0$, equations (1.23) yield

$$0 = i\omega Q_1 - i\omega Q_2,$$

$$Q_0 = Q_1 + Q_2 .$$

Solving for Q_1 and Q_2 we obtain

$$Q_1 = Q_2 = \frac{1}{2} Q_0 ,$$

which when substituted in (1.23) give

$$I = -\omega Q_0 \left\{ \frac{e^{i\omega t} - e^{-i\omega t}}{2i} \right\},$$

(1.25)

$$Q = Q_0 \left\{ \frac{e^{i\omega t} + e^{-i\omega t}}{2} \right\}.$$

We recognize the bracketed functions to be the sine and cosine of ωt, so that

$$I = -\omega Q_0 \sin\omega t = \omega Q_0 \cos(\omega t + \frac{\pi}{2}) ,$$

(1.26)

$$Q = Q_0 \cos\omega t .$$

From Theorem 3.1 of Chapter 1 we are assured that these two functions I and Q are the unique pair of functions satisfying the two simultaneous differential equations (1.5) and the initial conditions $Q(0) = Q_0$, $I(0) = 0$.

Had we solved instead the mechanical system with initial conditions $u(0) = U_0$, $v(0) = 0$ we can see from the analogies previously established that $\omega = \sqrt{\frac{k}{m}}$, and the eigenvalues $\lambda_1 = i\omega$, $\lambda_2 = -i\omega$ lead to two eigenvectors. The sum of solutions is again a solution and the results corresponding to (1.26) are

$$v = \omega U_0 \cos(\omega t + \tfrac{\pi}{2}) \ ,$$

(1.27)

$$u = U_0 \cos \omega t \ .$$

The nature of the response is the same in both systems; it is seen to be a continuing oscillation whose amplitude is given by the initial disturbance and whose frequency $f = \frac{\omega}{2\pi}$ is called the natural frequency. The velocity or current leads the displacement or charge by a phase angle of $\frac{\pi}{2}$.

The method of solution of two simultaneous first order linear differential equations can be summarized as follows. Assume a solution for each dependent variable in the form of a constant multiple of $e^{\lambda t}$. The differential equations become algebraic equations which when written in matrix form become the eigenvalue problem with λ as the eigenvalue. If the characteristic equation yields distinct eigenvalues λ_1 and λ_2, these give rise to two independent solutions. When these solutions are added together and the arbitrary constants evaluated from the initial conditions we obtain the unique solution to the original problem. Should the eigenvalues not be distinct, $\lambda_1 = \lambda_2$, then we do not obtain two independent solutions and the method outlined here does not apply. We will consider the case of multiple roots later in this chapter.

Example 1.2. Solve the system of differential equations

$$\frac{dx_1}{dt} = 2x_1 + a_{12}x_2 \,, \qquad \frac{dx_2}{dt} = x_1 \,, \tag{1.28}$$

subject to the initial conditions

$$x_1(0) = 2 \,, \; x_2(0) = 2. \tag{1.29}$$

Assume a solution of the form

$$x_1 = Be^{\lambda t} \,, \; x_2 = Ce^{\lambda t}. \tag{1.30}$$

Substitution in (1.28) gives, after dividing out $e^{\lambda t}$,

$$\lambda B = 2B + a_{12}C \,, \qquad \lambda C = B, \qquad \text{or } (\lambda I - A)\begin{pmatrix} B \\ C \end{pmatrix} = 0, \tag{1.31}$$

where A is the matrix of coefficients on the right hand side of (1.28) or (1.31).

Equations (1.31) have a solution if and only if the characteristic equation

$$|\lambda I - A| = \lambda^2 - 2\lambda - a_{12} = 0 \tag{1.32}$$

is satisfied. The values of λ are determined from

$$\lambda = \frac{1}{2}\left[2 \pm \sqrt{4 + 4a_{12}}\right] = 1 \pm \sqrt{1 + a_{12}}. \tag{1.33}$$

From (1.33) the eigenvalues λ are seen to be real, complex, or non-distinct depending on the value of a_{12}.

If $a_{12} > -1$ then the values of λ are real and distinct. For example, if $a_{12} = 3$ then we have

$$\lambda_1 = 3 \ , \ \lambda_2 = -1.$$

Either of equations (1.31) gives the relation between B and C corresponding to each eigenvalue. Thus, for λ_1 we have the solution

$$x_1 = 3C_1 e^{3t} \ , \ x_2 = C_1 e^{3t} \ , \tag{1.34}$$

and for λ_2 the solution is

$$x_1 = -C_2 e^{-t} \ , \ x_2 = C_2 e^{-t}. \tag{1.35}$$

By substitution in the original equations (1.28) it can be verified that the sum of solutions (1.34) and (1.35),

$$\begin{aligned} x_1 &= 3C_1 e^{3t} - C_2 e^{-t}, \\ x_2 &= C_1 e^{3t} + C_2 e^{-t} \ , \end{aligned} \tag{1.36}$$

is also a solution. Initial conditions (1.29) when used in (1.36) yield

$$\left\{ \begin{aligned} 2 &= 3C_1 - C_2 \ , \\ 2 &= C_1 + C_2 \ , \end{aligned} \right. \qquad \text{or } C_1 = C_2 = 1.$$

If $a_{12} = -1$ then from (1.33) the eigenvalues λ are not distinct and we do not have two solutions of the form (1.30). The solution in this case will be considered later.

If $a_{12} < -1$ then the square root in (1.33) is imaginary and the eigenvalues are complex numbers. For example, say $a_{12} = -2$ and we have

$$\lambda_1 = 1 + i \ , \ \lambda_2 = 1 - i .$$

Again (1.31) gives the relation between B and C corresponding to each eigenvalue. For λ_1 the solution is

$$x_1 = (1+i)C_1 e^{(1+i)t} \ , \ x_2 = C_1 e^{(1+i)t} \ , \tag{1.37}$$

while for λ_2 the solution is

$$x_1 = (1-i)C_2 e^{(1-i)t} \ , \ x_2 = C_2 e^{(1-i)t} . \tag{1.38}$$

The sum of solutions (1.37) and (1.38) again satisfies (1.28) so that

$$\begin{aligned} x_1 &= (1+i)C_1 e^{(1+i)t} + (1-i)C_2 e^{(1-i)t} \ , \\ x_2 &= C_1 e^{(1+i)t} + C_2 e^{(1-i)t} \ , \end{aligned} \tag{1.39}$$

is also a solution. Initial conditions (1.29) when used in (1.39) yield

$$\left\{ \begin{aligned} 2 &= (1+i)C_1 + (1-i)C_2 \ , \\ 2 &= C_1 + C_2 \ , \end{aligned} \right. \qquad \text{or } C_1 = C_2 = 1. \tag{1.40}$$

The final simplification of the solution is left to the reader (Problem 1.2).

Problems

1.1 Find the condition on a_{11} for the system of equations

$$\frac{dx_1}{dt} = a_{11}x_1 - 2x_2 ,$$

$$\frac{dx_2}{dt} = x_1 + 2x_2 ,$$

to have real eigenvalues. Find the solutions corresponding to each eigenvalue and the sum of solutions when $a_{11} = 5$. Partial Answer: $a_{11} \geqq 2 + 2\sqrt{2}$ or $a_{11} \leqq 2 - 2\sqrt{2}$.

1.2 In Example 1.2 show that when the constants from (1.40) are substituted in solution (1.39) that solution can be written in the form

$$x_1 = 2e^t \{\cos t - \sin t\} ,$$

$$x_2 = 2e^t \cos t.$$

1.3 If in Example 1.1 the circuit elements are not ideal then a series resistance R must be added to the circuit of Fig. 1.2. Derive the system of equations analogous to (1.5) and verify that the characteristic equation is

$$\lambda^2 + \frac{R}{L}\lambda + \frac{1}{LC} = 0.$$

$\frac{1}{LC}$ and find the solution for the two cases:

when $\left(\frac{R}{2L}\right)^2 > \frac{1}{LC}$ and when $\left(\frac{R}{2L}\right)^2 < \frac{1}{LC}$.

1.4 For the system of equations

$$\frac{dx_1}{dt} = a_{11}x_1 - 2x_2 \ ,$$

$$\frac{dx_2}{dt} = x_1 + a_{22}x_2 \ ,$$

find the relation which must exist between a_{11} and a_{22} in order that at least one of the eigenvalues is zero. If in the system of equations $a_{11} = 1$ and $a_{22} = -2$ find the solution which satisfies the initial conditions $x_1(0) = 3$, $x_2(0) = 1$.

Partial Answer: $x_1 = 4 - e^{-t}$, $x_2 = 2 - e^{-t}$.

1.5 Figure 1.3 shows a rigid disk, with moment of inertia I (slug-in.^{2}) about the vertical central axis, fastened to a thin tube of negligible mass. The torsional stiffness of the tube is k $\left(\frac{\text{in.-lb.}}{\text{radian}}\right)$.

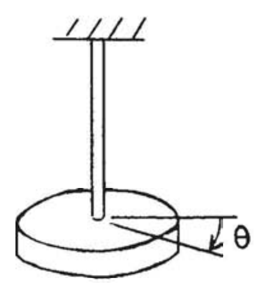


Figure 1.3

The disk is rotated through an initial angle θ_0 and released from rest. Using the definition of angular momentum, H, as $H = I\dot{\theta}$, and the principle that the time rate of change of angular momentum equals the torque, derive the system of equations analogous to (1.5) and find their solution. Establish the analogous quantities with the electrical and mechanical systems of Example 1.1.

1.6 Find the solution to the system of equations

$$\frac{dx}{dt} = -3x + 2y \ , \qquad \frac{dy}{dt} = -2y \ ,$$

subject to the initial conditions $x(0) = 1$, $y(0) = 1$. First use the method of this chapter and then check the result by using the direct approach of Chapter 1 after observing that the second equation is uncoupled from the first.

Answer: $x = 2e^{-2t} - e^{-3t}$, $y = e^{-2t}$.

1.7 An inverted pendulum is shown in Figure 1.4 Assume the mass M to be concentrated at the end point of the rod. The ideal, linear spring has stiffness k. Gravity acts downwards as shown. The pendulum is rotated from its equilibrium position $\theta = 0$ through an initial angle θ_0 and released from rest. As in Problem 1.5 use the definition of H and the torque principle with respect to the point 0 to show that for initial angles such that $\sin\theta \doteq \theta$ the motion is described by the system of equations

$$\frac{dH}{dt} = (MgL - kL^2)\theta \ ,$$

$$ML^2 \ \frac{d\theta}{dt} = H.$$

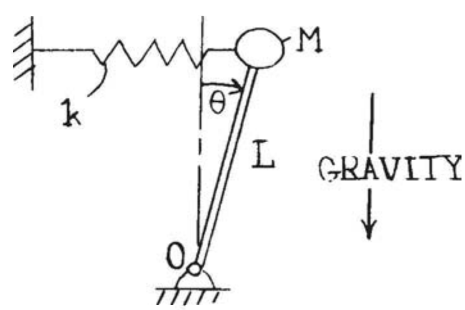


Figure 1.4

Find the solution to this system of equations. Distinguish between the two cases when $kL > Mg$ and when $kL < Mg$.

1.8 Consider the system of equations

$$\frac{dx}{dt} + \frac{dy}{dt} = -2y ,$$

$$\frac{dx}{dt} - \frac{dy}{dt} = -x.$$

Put these equations into the form (1.1) and show that the characteristic equation is

$$2\lambda^2 + 3\lambda + 2 = 0.$$

Now, rather than eliminating one derivative in each of the given equations, assume directly that their solution is of exponential form

$$x = Xe^{\lambda t} , \; y = Ye^{\lambda t} ,$$

where X and Y are constants. Substitute this solution in the given equations and show that the same characteristic equation is obtained.

Generalize this result to show that the solution of the system of equations

$$b_{11} \frac{dx_1}{dt} + b_{12} \frac{dx_2}{dt} = a_{11}x_1 + a_{12}x_2 ,$$

$$b_{21} \frac{dx_1}{dt} + b_{22} \frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 ,$$

leads to the solution of the homogeneous set of equations

$$(B\lambda - A) x = 0 ,$$

where B is the matrix of coefficients b_{ij}, A the matrix of coefficients a_{ij} and x the vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

2. <u>Systems of First Order Homogeneous Linear Differential Equations with Constant Coefficients</u>. <u>General Case</u>.

In Section 1 we considered examples of special cases of the system of two equations

$$(2.1) \qquad \begin{cases} \dfrac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 , \\ \dfrac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 , \end{cases} \qquad \text{or} \qquad \frac{dx}{dt} = Ax ,$$

with initial conditions

$$(2.2) \qquad x_1(t_0) = X_{10} , \; x_2(t_0) = X_{20}.$$

In order to solve equations (2.1) assume a solution of the form

$$(2.3) \qquad x_1 = X_1 e^{\lambda t} , \; x_2 = X_2 e^{\lambda t} ,$$

in which X_1, X_2 and λ are constants to be determined. For (2.3) to be a solution, it must satisfy (2.1). Hence, substitution of (2.3) into (2.1) yields the condition

$$(2.4) \qquad (\lambda I - A)X = 0.$$

Equation (2.4) is the matrix equation for the eigenvalues λ of the matrix of coefficients $A = (a_{ij})$. Thus, X is the eigenvector having constant components X_1, X_2. We know that (2.4) has a solution if and only if $\det(\lambda I - A)$ is zero, so that the characteristic equation is

$$(2.5)\qquad \begin{vmatrix} \lambda-a_{11} & -a_{12} \\ -a_{21} & \lambda-a_{22} \end{vmatrix} = \lambda^2-(a_{11}+a_{22})\lambda + (a_{11}a_{22}-a_{12}a_{21}) = 0.$$

Assume that (2.5) determines two distinct eigenvalues λ_1 and λ_2. Then with eigenvalue λ_1, from (2.3) we have a solution

$$(2.6)\qquad x_1 = X_{11}e^{\lambda_1 t}, \quad x_2 = X_{21}e^{\lambda_1 t}.$$

Constants X_{11} and X_{21} are associated with eigenvalue λ_1, and (2.4) gives the following relation between X_{11} and X_{21},

$$(2.7)\qquad (\lambda_1 I-A)\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} = 0.$$

only in letters alph. X, A

When written out, equation (2.7) gives either

$$(2.8)\quad (\lambda_1-a_{11})X_{11}-a_{12}X_{21} = 0 \quad\text{or}\quad -a_{21}X_{11}+(\lambda_1-a_{22})X_{21} = 0.$$

These two equations necessarily are satisfied by the same ratio of X_{11} to X_{21}. A second solution comes from eigenvalue λ_2 so that (2.3) gives

$$(2.9)\qquad x_1 = X_{12}e^{\lambda_2 t}, \quad x_2 = X_{22}e^{\lambda_2 t}$$

where constants X_{12} and X_{22} are associated with λ_2. Equation (2.4) gives the following relation between X_{12} and X_{22},

$$(2.10) \qquad (\lambda_2 I - A)\begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix} = 0 ,$$

which when written out gives either

$$(2.11) \quad (\lambda_2 - a_{11})X_{12} - a_{12}X_{22} = 0 \text{ or } -a_{21}X_{12} + (\lambda_2 - a_{22})X_{22} = 0.$$

As in the examples of the previous section, we add the two solutions associated with each distinct eigenvalue to obtain

$$(2.12) \qquad \begin{aligned} x_1 &= X_{11}e^{\lambda_1 t} + X_{12}e^{\lambda_2 t} , \\ x_2 &= X_{21}e^{\lambda_1 t} + X_{22}e^{\lambda_2 t} . \end{aligned}$$

Direct substitution of this sum of solutions, equations (2.12), into the original system of differential equations (2.1) shows that the original equations are satisfied. Hence, equations (2.12) are a solution and contain two arbitrary constants, say X_{11} and X_{12}, by which initial conditions (2.2) may be satisfied. Thus, for the case of distinct eigenvalues of A we have a general method to construct the solution.

It should be clear from this work on two simultaneous equations and the previous considerations in Chapter 2 on systems of linear equations, that it should be possible to analyze the general case of n differential equations in n unknowns. Thus, consider the system of equations

$$\frac{dx_1}{dt} = a_{11}x_1 + \cdot \cdot \cdot + a_{1n}x_n ,$$

$$\frac{dx_2}{dt} = a_{21}x_1 + \cdot \cdot \cdot + a_{2n}x_n ,$$

(2.13)

$$\cdot \cdot \cdot$$

$$\frac{dx_n}{dt} = a_{n1}x_1 + \cdot \cdot \cdot + a_{nn}x_n ,$$

which can still be written, as in (2.1), in the form

$$\frac{dx}{dt} = Ax. \qquad (2.14)$$

The following theorem forms the basis of the whole mathematical theory underlying the solution of equation (2.14).

Theorem 2.1. The set W of all solutions of equation (2.14) is an n-dimensional vector space.

Proof. A solution of (2.14) is a set of n functions of t; that is,

$$x(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{pmatrix} ,$$

where each $f_i(t)$ is an ordinary scalar function of t, as were x_1 and x_2 of equations (2.3) in the case of two variables. To see that the set of all such solutions forms a vector space we must refer back to the definition of a vector space (Definition 3.1 in Chapter 2) and see that all the conditions are satisfied. First we check closure under addition and multiplication by scalars. If x and y are two solutions of (2.14) and we let $z = x + y$ then

$$\frac{dz}{dt} = \frac{d(x+y)}{dt} = \frac{dx}{dt} + \frac{dy}{dt}$$

$$= Ax + Ay = A(x+y) = Az ,$$

and so z is also a solution of (2.14). Closure under multiplication by scalars is checked similarly. The vector 0, which is the n-tuple of functions each of which is identically zero, is obviously a solution since $\frac{d0}{dt} = 0$ and $A0 = 0$. Finally, the eight identities of Definition 3.1 are easily seen to hold for n-tuples of functions just as for n-tuples of numbers. <u>Thus W is a vector space</u>.

We still must prove that the dimension of W is n. This will be done by constructing a basis for W consisting of n solutions. These n solutions are assumed known from the fundamental existence theorem stated in Chapter 1 (Theorem 3.1). For (2.14) in terms of components is (2.13) and the partial derivatives $\frac{\partial f_j}{\partial y_k}$ that appear in the existence theorem are merely our constants a_{jk} that form matrix A. Being constants, they are certainly continuous in any domain D and the existence theorem can be applied. In our case the initial conditions take the form that given any n-tuple of constants, v_0, there exists a unique solution $x(t)$ such that $x(t_0) = v_0$.

To construct a basis for W we first get a basis for the space V_n of n-tuples. The simplest such basis is $\{e_1,\cdots,e_n\}$, where e_i has its i-th component equal to 1 and all other components are zero.

Let* $x^i(t)$ be the solution with initial condition $x^i(t_0) = e_i$. We shall prove that $\{x^1(t), x^2(t), \cdots x^n(t)\}$ is a basis for W and hence W is n-dimensional. First we prove that the $x^i(t)$ are independent. Suppose we had a linear relation

$$\sum_{i=1}^{n} c_i x^i(t) = 0$$

the c_i being constants. Then for the special value $t = t_0$,

$$\sum_{i=1}^{n} c_i x^i(t_0) = \sum_{i=1}^{n} c_i e_i = 0. \tag{2.15}$$

But $\sum_{i=1}^{n} c_i e_i = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$, and so (2.15) implies that all the c_i are zero. Hence, the $x^i(t)$ satisfy only the trivial relation and so are independent (Theorem 6.1 of Chapter 2).

Now finally, we must prove that $\{x^1(t), \cdots, x^n(t)\}$ spans W. That is, if $x(t)$ is <u>any</u> solution of (2.14) then $x(t)$ is a linear combination of $\{x^1(t), \cdots, x^n(t)\}$. Let

$$x(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix} \tag{2.16}$$

by any solution of (2.14). Now the vector formed by the linear combination of solutions $x^i(t)$,

$$\sum_{i=1}^{n} f_i(t_0) x^i(t) , \tag{2.17}$$

* We often refer to several different vectors that are solutions of (2.14). We cannot distinguish these by subscripts, since subscripts are used for components of a vector. So we use <u>superscripts</u>, as in $x^1(t)$, $z^1(t)$, y^2, etc. It must be understood that these are not exponents <u>but</u> <u>merely</u> distinguishing indices.

is also a solution of (2.14), for we have already proved that the set of solutions is a vector space. Further, at $t = t_0$ the linear combination (2.17) becomes

$$\sum_{i=1}^{n} f_i(t_0)x^i(t_0) = \sum_{i=1}^{n} f_i(t_0)e_i = \begin{pmatrix} f_1(t_0) \\ \vdots \\ f_n(t_0) \end{pmatrix} ,$$

and hence equals $x(t_0)$ from (2.16). Consequently, solutions (2.17) and (2.16) are equal at $t = t_0$ and by the uniqueness theorem it follows that they must be identical for all t. Hence

$$x(t) = \sum_{i=1}^{n} c_i x^i(t) \ , \ c_i = f_i(t_0) \ , \tag{2.18}$$

and $\{x^1(t),\ldots,x^n(t)\}$ spans W. This completes the proof of Theorem 2.1.

<u>In summary</u>: the set of all solutions of (2.14) is an n-dimensional vector space, and <u>any</u> solution can be written as a linear combination of a set of n independent solutions. It is worth noting that nowhere in this proof have we made any essential use of the fact that the elements a_{ij} of A are constants. The proof will still work if the a_{ij} are any functions of the independent variable t which are continuous over some range of values $a \leqq t \leqq b$ containing t_0. Thus, the important point is that the right-hand side of (2.14) is a homogeneous linear function of the dependent variables, and not the fact that the coefficients are constants.

Theorem 2.1 shows why in the case of two simultaneous equations, the examples of Section 1 and the method given at

the beginning of this section, one is led to the eigenvalue problem with vector solutions which span a 2-dimensional space and hence can be added to form a solution. Theorem 2.1 does not however give any indication as to how one finds n independent vector solutions $\{x^1(t), x^2(t), \ldots, x^n(t)\}$ of (2.14), for their existence was assumed on the basis of the existence theorem of Chapter 1. To find such solutions for the n-dimensional case, we follow the same method as in the 2-dimensional case, and look for solutions of the form

$$x(t) = ve^{\lambda t} , \tag{2.19}$$

where v is a constant non-zero vector and λ a scalar. Substituting (2.19) into (2.14) gives

$$\lambda ve^{\lambda t} = Ave^{\lambda t}.$$

This is an identity if and only if

$$\lambda v = Av.$$

Hence λ must be an eigenvalue of A and v an associated eigenvector. With Theorem 2.1 in mind, we now ask whether we can find n independent solutions of the form (2.19). The answer depends entirely on the eigenvectors v. If we can find n independent eigenvectors of A then the corresponding solutions of the form (2.19) will also be independent, for we have $v = x(0)$ and if the x(t) are independent for $t = 0$ they are certainly independent as functions of t. We have seen in Chapter 5, Theorem 5.6, that any symmetric matrix has n independent eigenvectors. Hence if A is symmetric we can get a complete solution of (2.14) by linear combinations of the type (2.19). Another case that gives n independent solutions arises when the n eigenvalues are all different (Chapter 5, Theorem 7.3). These two cases occur often enough to make the method outlined here of great importance.

On the other hand it can be shown that if no independent set of n eigenvectors exists then solutions of a form different from (2.19) must be considered; we shall not investigate this situation in general but will consider an important special case in Section 6. See also Problems 2.9-2.12. A discussion of the general case can be found in Coddington and Levinson, Theory of Ordinary Differential Equations, McGraw-Hill Book Company, New York, 1955, and in W. Kaplan, Ordinary Differential Equations, Addison-Wesley Publishing Company, Reading, Mass., 1958.

Example 2.1. Solve the system of equations

$$(2.20)\qquad \begin{cases} \dfrac{dz_1}{dt} = 4z_1 - 2z_2 , \\ \\ \dfrac{dz_2}{dt} = z_1 + z_2 , \end{cases} \quad \text{or } \frac{dz}{dt} = Az ,$$

subject to the initial conditions

$$(2.21)\qquad z_1(0) = 3 \quad , \quad z_2(0) = 2 \; ;$$

and find the base vectors $x^1(t)$, $x^2(t)$ of Theorem 2.1.

To find the solutions of the given equations assume

$$z = ve^{\lambda t} ;$$

substitution yields the relation

$$(\lambda I - A)v = 0 , \tag{2.22}$$

which has a solution v if and only if

$$\det(\lambda I - A) = \begin{vmatrix} \lambda-4 & 2 \\ -1 & \lambda-1 \end{vmatrix} = \lambda^2 - 5\lambda + 6 = 0 .$$

The roots of this characteristic equation give two different eigenvalues

$$\lambda_1 = 3 , \lambda_2 = 2$$

so that we are assured that two independent solutions can be obtained.

We have the corresponding two solutions

$$v_1 e^{3t} = \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} e^{3t} , \quad v_2 e^{2t} = \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix} e^{2t}.$$

Substitution of λ_1 and eigenvector $v_1 = \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix}$ into (2.22) gives

$$\begin{pmatrix} (3-4) & 2 \\ -1 & (3-1) \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = 0,$$

which when expanded gives

$$-v_{11} + 2v_{12} = 0.$$

A simple solution of this equation is $v_{11} = 2$, $v_{12} = 1$, and so we get one solution of (2.20) to be

$$z^1(t) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} e^{3t}.$$

In a similar manner, substitution of λ_2 and eigenvector $v_2 = \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix}$ into (2.22) yields

$$-v_{21} + v_{22} = 0,$$

which is satisfied by $v_{21} = 1$, $v_{22} = 1$, so that the second solution of (2.20) is

$$z^2(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{2t}.$$

These two solutions are independent since, in particular, the two vectors

$$z^1(0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ and } z^2(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

are independent.

From Theorem 2.1 we know that any solution can be written as a linear combination of these two solutions; thus

$$z(t) = c_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} e^{3t} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{2t}. \tag{2.23}$$

The two solutions $z^1(t)$ and $z^2(t)$ form a basis for the space W of all solutions, and for most purposes serve just as well as the special basis used in the proof of Theorem 2.1.

However, in this Example we have asked for this special basis so we proceed to construct it.

The solution $x^1(t)$ was defined by the initial condition $x^1(0) = e_1$, so that (2.23) gives

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = c_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2c_1 + c_2 \\ c_1 + c_2 \end{pmatrix},$$

from which we get $c_1 = 1$, $c_2 = -1$. Hence

$$x^1(t) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} e^{3t} + \begin{pmatrix} -1 \\ -1 \end{pmatrix} e^{2t}. \tag{2.24}$$

Similarly, $x^2(t)$ is such that $x_2(0) = e_2$. In this case (2.23) gives

$$e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2c_1 + c_2 \\ c_1 + c_2 \end{pmatrix}$$

and consequently $c_1 = -1$, $c_2 = 2$, and

$$x^2(t) = \begin{pmatrix} -2 \\ -1 \end{pmatrix} e^{3t} + \begin{pmatrix} 2 \\ 2 \end{pmatrix} e^{2t}. \tag{2.25}$$

The basis $\{x^1(t), x^2(t)\}$, while more complicated than the basis $\{z^1(t), z^2(t)\}$, is convenient if several different sets of initial conditions are to be considered. Consider the work involved in finding a solution of (2.20) with initial condition $z(0) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. If we use the z-basis we must solve the simultaneous equations

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} = c_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

for c_1 and c_2 to get the solution in the form (2.23). Using the x-basis we can express the solution in the form

$$z(t) = b_1 x^1(t) + b_2 x^2(t). \tag{2.26}$$

The corresponding condition on b_1 and b_2 is

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} = b_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

or, obviously, $b_1 = 3$, $b_2 = 2$. We can thus get the solution in the form (2.26) directly as

$$\begin{aligned} z(t) &= 3x^1(t) + 2x^2(t) \\ &= 3\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix} e^{3t} + \begin{pmatrix} -1 \\ -1 \end{pmatrix} e^{2t}\right] + 2\left[\begin{pmatrix} -2 \\ -1 \end{pmatrix} e^{3t} + \begin{pmatrix} 2 \\ 2 \end{pmatrix} e^{2t}\right] \\ &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} e^{3t} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{2t}. \end{aligned}$$

<u>Example 2.2</u>. **Relaxation of Resistor-Capacitor Networks.**

There are many physical situations in which an initially disturbed system requires a characteristic period of time to attain an equilibrium state. One example of such a system is the network of resistors and capacitors shown in Figure 2.1. Assume that initially the switch s in branch 1-2 is open, a battery of voltage V_0 is placed across the capacitor in vertical branch 0-1, and hence that capacitor acquires a charge CV_o, while the other capacitors are uncharged. The battery is then removed and at time $t = 0$

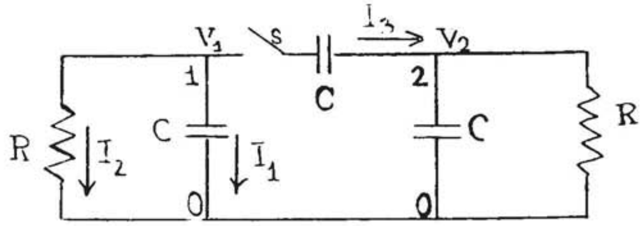


Figure 2.1

the switch is closed. The charged capacitor discharges through the network. The energy initially stored in the charged capacitor is dissipated in the resistors so that the equilibrium state has all capacitors discharged and no current flows.

In deriving the differential equations which describe such a network, one must choose either voltage, current, or charge as the variable. It is not always obvious which choice is most convenient and one must learn this by experience. As the unknown dependent variables, we will use the voltage from node 0 to 1, V_1, and the voltage from node 0 to 2, V_2. The assumed directions of current flow are shown on Figure 2.1. Due to voltage V_1 the current I_2 leaving node 1 flowing through the resistor at the left is

$$I_2 = \frac{1}{R} V_1. \tag{2.27}$$

Due to voltage V_1 the current I_1 flowing to capacitor C is obtained through the relation $Q_1 = CV_1$ so that

$$I_1 = \frac{dQ_1}{dt} = C \frac{dV_1}{dt} . \tag{2.28}$$

Similarly, the current I_3 leaving node 1 in branch 1-2 due to the potential difference $V_1 - V_2$ is

$$I_3 = C \frac{d(V_1 - V_2)}{dt}. \tag{2.29}$$

Using Kirchhoff's law of conservation of charge at node 1 we conclude that the sum of these three currents must be zero:

$$\frac{1}{R} V_1 + C \frac{dV_1}{dt} + C \frac{d}{dt} (V_1 - V_2) = 0. \tag{2.30}$$

In a similar manner, the conservation of charge at node 2 gives

$$\frac{1}{R} V_2 + C \frac{dV_2}{dt} - C \frac{d}{dt} (V_1 - V_2) = 0. \tag{2.31}$$

Equations (2.30) and (2.31) together with initial conditions

$$V_1(0) = V_0 \quad , \quad V_2(0) = 0, \tag{2.32}$$

are the governing equations. Equations (2.30) and (2.31) can be put in the form

$$\begin{aligned} 2C \frac{dV_1}{dt} - C \frac{dV_2}{dt} &= - \frac{1}{R} V_1 , \\ -C \frac{dV_1}{dt} + 2C \frac{dV_2}{dt} &= - \frac{1}{R} V_2 . \end{aligned} \tag{2.33}$$

Although it is possible to solve equations (2.33) directly by assuming a solution of exponential form (see Problem 1.8), we will put (2.33) in the standard form of (2.13) or (2.14). If the first equation of (2.33) is multiplied by 2 and added to the second equation, $\frac{dV_2}{dt}$ is eliminated. Similarly, multiplication of the second equation by 2 and addition with the first equation eliminates $\frac{dV_1}{dt}$.

Thus, in standard form (2.33) becomes

$$\begin{aligned} \frac{dV_1}{dt} &= - \frac{2}{3RC} V_1 - \frac{1}{3RC} V_2 , \\ \frac{dV_2}{dt} &= - \frac{1}{3RC} V_1 - \frac{2}{3RC} V_2 . \end{aligned} \tag{2.34}$$

Following the method previously used, we seek a solution of the form

(2.35) $$V_1 = E_1 e^{\lambda t} \ , \ V_2 = E_2 e^{\lambda t} \ ,$$

so that upon substitution in (2.34) we obtain

(2.36) $$\lambda E = AE \ ,$$

where

(2.37) $$A = \begin{pmatrix} -\frac{2}{3RC} & \frac{-1}{3RC} \\ -\frac{1}{3RC} & \frac{-2}{3RC} \end{pmatrix} \ , \ E = \begin{pmatrix} E_1 \\ E_2 \end{pmatrix} .$$

Matrix A is symmetric, consequently we know that the eigenvalues are real. The characteristic equation $\det(\lambda I - A) = 0$ is

(2.38) $$\left(\lambda + \frac{2}{3RC}\right)^2 - \left(\frac{1}{3RC}\right)^2 = 0,$$

from which we obtain the eigenvalues

(2.39) $$\lambda_1 = -\frac{1}{3RC} \ , \ \lambda_2 = -\frac{1}{RC} \ .$$

Equation (2.36) with eigenvalue λ_1 gives the relation $E_{21} = -E_{11}$ so that one solution is

(2.40) $$V(t) = \begin{pmatrix} E_{11} \\ -E_{11} \end{pmatrix} e^{-\frac{1}{3RC}t} .$$

With eigenvalue λ_2, equation (2.36) gives the relation $E_{22} = E_{12}$ so that a second solution is

$$V(t) = \begin{pmatrix} E_{12} \\ E_{12} \end{pmatrix} e^{-\frac{1}{RC} t} . \tag{2.41}$$

If a relaxation time is the time required for the voltages to fall to 1/e of their previous values, then we see that solution (2.40) associated with λ_1 gives rise to the "slow" relaxation time $t_1 = 3RC$ while solution (2.41) associated with λ_2 gives rise to the "fast" relaxation time $t_2 = RC$. Solution (2.41) shows that in the "fast" configuration $V_1 = V_2$ so that there is no voltage drop $(V_1 - V_2)$ across the coupling capacitor C in branch 1-2 and hence no current in this branch.

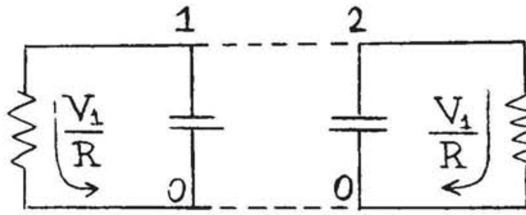


Figure 2.2

The currents in the two end loops are equal to $\frac{V_1}{R}$ and are shown in Figure 2.2. In this configuration the original differential equations (2.30) and (2.31) are uncoupled, and we see that physically the two end loops are uncoupled.

In the "slow" configuration solution, (2.40) shows that $V_1 = -V_2$, and now there is a voltage across branch 1-2. The currents are easily obtained from (2.27), (2.28), and (2.29), and are shown in Figure 2.3. Of course, in both configurations the

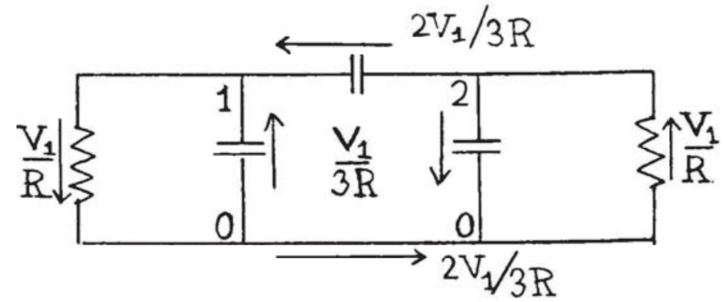


Figure 2.3

voltages and currents decay with time, but each configuration has its characteristic relaxation time.

We know that any solution is the sum of solutions (2.40) and (2.41) so that

$$(2.42) \qquad V(t) = \begin{pmatrix} E_{11} \\ -E_{11} \end{pmatrix} e^{-\frac{1}{3RC}t} + \begin{pmatrix} E_{12} \\ E_{12} \end{pmatrix} e^{-\frac{1}{RC}t}.$$

Constants E_{11} and E_{12} are determined by initial conditions (2.32). However, observe that regardless of the particular initial conditions the voltages in the network are the result of the superposition of the two configurations corresponding to λ_1 and λ_2. A particular case is determined by merely taking a certain fraction of the "slow" configuration and a certain fraction of the "fast" configuration. With the initial conditions of (2.32) we obtain

$$(2.43) \qquad V(0) = \begin{pmatrix} V_0 \\ 0 \end{pmatrix} = \begin{pmatrix} E_{11} \\ -E_{11} \end{pmatrix} + \begin{pmatrix} E_{12} \\ E_{12} \end{pmatrix},$$

from which we get $E_{11} = E_{12} = V_0/2$ so that half the initial voltage is associated with each configuration. Substitution in (2.42) yields the final solution

$$(2.44) \qquad V(t) = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} V_0/2 \\ -V_0/2 \end{pmatrix} e^{-\frac{1}{3RC}t} + \begin{pmatrix} V_0/2 \\ V_0/2 \end{pmatrix} e^{-\frac{1}{RC}t}.$$

Example 2.3. Relaxation of Resistor-Inductor Networks

The three loop network of Figure 2.4 when initially disturbed will tend to an equilibrium state of no voltages or currents because resistors R remove energy from the system. We wish to find expressions for the three loop currents I_1, I_2, and I_3.

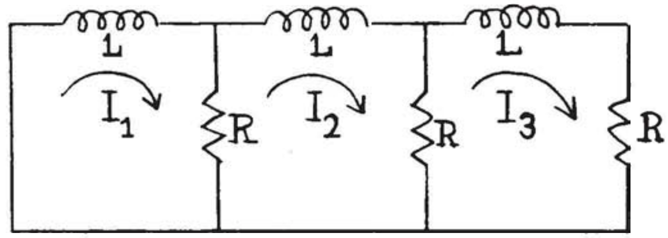


Figure 2.4

We use the three currents as the unknown variables and application of Kirchhoff's voltage law around each loop leads to the three equations

$$L\frac{dI_1}{dt} + R(I_1-I_2) = 0 ,$$

$$(2.45) \qquad L\frac{dI_2}{dt} + R(I_2-I_3) - R(I_1-I_2) = 0 ,$$

$$L\frac{dI_3}{dt} + RI_3 - R(I_2-I_3) = 0.$$

These equations when in the form (2.13) are

$$(2.46) \quad \begin{cases} \dfrac{dI_1}{dt} = -\dfrac{R}{L}I_1 + \dfrac{R}{L}I_2 , \\[2ex] \dfrac{dI_2}{dt} = +\dfrac{R}{L}I_1 - \dfrac{2R}{L}I_2 + \dfrac{R}{L}I_3 , \qquad \text{or} \qquad \dfrac{dI}{dt} = AI , \\[2ex] \dfrac{dI_3}{dt} = \qquad\qquad + \dfrac{R}{L}I_2 - \dfrac{2R}{L}I_3 , \end{cases}$$

where

$$A = \frac{R}{L}\begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix} , \qquad I = \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix} . \tag{2.47}$$

Since A is symmetric we know that it has three real eigenvalues, and the eigenvectors will be independent (and orthogonal, but we make no use of this). The eigenvalues of A are just R/L times those of the matrix

$$A_1 = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix} , \tag{2.48}$$

which are the roots of $\det(\lambda I - A_1) = 0$:

$$\begin{vmatrix} \lambda+1 & -1 & 0 \\ -1 & \lambda+2 & -1 \\ 0 & -1 & \lambda+2 \end{vmatrix} = \lambda^3 + 5\lambda^2 + 6\lambda + 1 = 0. \tag{2.49}$$

The three roots of (2.49), to 2 decimal places, are

$$\lambda_1 = -0.20 , \ \lambda_2 = -1.56 , \ \lambda_3 = -3.25 , \tag{2.50}$$

and associated eigenvectors are

$$\begin{pmatrix} 2.35 \\ 1.80 \\ 1.00 \end{pmatrix} , \begin{pmatrix} -0.79 \\ -0.44 \\ 1.00 \end{pmatrix} , \begin{pmatrix} 0.56 \\ -1.25 \\ 1.00 \end{pmatrix} . \tag{2.51}$$

We construct any solution by taking a linear combination of these three independent solutions. Hence

$$(2.52) \quad \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix} = c_1 \begin{pmatrix} 2.35 \\ 1.80 \\ 1.00 \end{pmatrix} e^{-.20 \frac{R}{L} t} + c_2 \begin{pmatrix} -0.79 \\ -0.44 \\ 1.00 \end{pmatrix} e^{-1.56 \frac{R}{L} t}$$

$$+ c_3 \begin{pmatrix} 0.56 \\ -1.25 \\ 1.00 \end{pmatrix} e^{-3.25 \frac{R}{L} t}$$

As in Example 2.2 we see that the solution is a superposition of solutions, each having a cnaracteristic relaxation time. This case lacks the symmetry of Example 2.2 so that the current components of each eigenvector bear no simple relation one to the other. However, each eigenvector defines a current configuration associated with each relaxation time. Constants c_1, c_2, c_3 would be found from initial conditions $I_1(0)$, $I_2(0)$, and $I_3(0)$.

Problems

2.1 Consider the system of equations

$$\frac{dx}{dt} = 2x - y ,$$

$$\frac{dy}{dt} = -3x ,$$

with initial conditions $x(0) = X_0$, $y(0) = Y_0$.

(a) Find the eigenvalues and associated eigenvectors obtained in the process of solving the given equations. Partial answer: $\lambda_1 = 3$, $\lambda_2 = -1$.

(b) Determine the base vectors $x^1(t)$, $x^2(t)$ which appear in the proof of Theorem 2.1. Then find the solution which satisfies the initial conditions.

Partial Answer: $x^2(t) = \begin{pmatrix} -1/4 \\ 1/4 \end{pmatrix} e^{3t} + \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix} e^{-t}.$

(c) What relation must exist between initial conditions X_0 and Y_0 so that the e^{-t} term does not appear in the solution?

2.2 Consider the system of equations

$$\frac{dx_1}{dt} = x_1 - 4x_2 + x_3 ,$$

$$\frac{dx_2}{dt} = \quad - 2x_2 + x_3 ,$$

$$\frac{dx_3}{dt} = \qquad 4x_3 .$$

(a) Find the eigenvalues and associated eigenvectors.

Partial Answer: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3/4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3/2 \\ 9 \end{pmatrix}$.

(b) Determine the base vectors $x^1(t)$, $x^2(t)$, $x^3(t)$ which appear in the proof of Theorem 2.1.

Partial Answer: $x^2(t) = \begin{pmatrix} -4/3 \\ 0 \\ 0 \end{pmatrix} e^{t} + \begin{pmatrix} 4/3 \\ 1 \\ 0 \end{pmatrix} e^{-2t}.$

2.3 Find the solution of the system of differential equations

$$\frac{dx}{dt} = x + 3y \ ,$$

$$\frac{dy}{dt} = x - y \ ,$$

$$\frac{dz}{dt} = -\ 6y + 4z \ ,$$

which satisfies the initial conditions $x(0) = 2$, $y(0) = 2$, $z(0) = -1$.

Partial Answer: $y = e^{2t} + e^{-2t}$.

2.4 Consider the system of equations

$$\frac{dx_1}{dt} = -2x_1 + 2x_2 \ ,$$

$$\frac{dx_2}{dt} = -\ x_1 - 4x_2 \ ,$$

with initial conditions $x_1(0) = 0$, $x_2(0) = 1$.

(a) Find the eigenvectors associated with each eigenvalue.

Answer: $\begin{pmatrix} 1 \\ (-1+i)/2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -(1+i)/2 \end{pmatrix}$.

(b) Find the solution of the given equations with the given initial conditions.

Partial Answer: $x_2 = \frac{1}{2i}\left\{ (-1+i)e^{(-3+i)t} + (1+i)e^{-(3+i)t} \right\}$.

(c) Write the solution of part (b) in terms of real quantities.

Partial Answer: $x_2 = e^{-3t}\{\cos t - \sin t\}$.

2.5 Consider the network of capacitors and resistors shown in Figure 2.5. Originally the switch is on contact s_1 and a battery of voltage V_0 is connected across the $\frac{1}{4}C$ capacitor. At time $t = 0$ the switch is flipped to contact s_2 in negligible time and the battery removed.

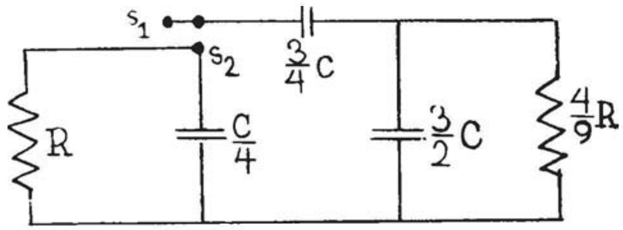


Figure 2.5

(a) Determine the system of node voltage equations analogous to (2.30),(2.31) and (2.34).

(b) Show that the eigenvalues are -2/3RC and -2/RC and determine the associated eigenvectors. Hence, identify the two relaxation configurations.

(c) Find the solution which satisfies the initial conditions.

2.6 Torsional oscillations in shafts are of major importance in the design of rotating machinery, for torsional vibrations and resonance can occur. Figure 2.6 shows two flywheels of moment of inertia I_1 and I_2 attached to a shaft having negligible mass which is free to rotate in bearings at its ends. The torsional stiffness of the shaft between

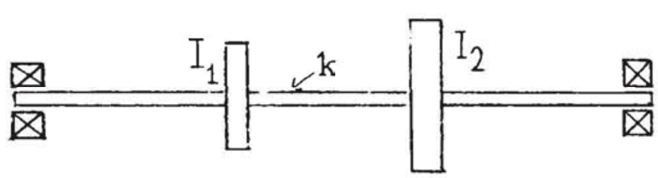


Figure 2.6

the two disks is k, so that the torque on each is k times the relative angular displacement.

(a) Show that the system of equations governing the motion is

$$\frac{dH_1}{dt} = -k(\theta_1-\theta_2) , \quad I_1 \frac{d\theta_1}{dt} = H_1 ,$$

$$\frac{dH_2}{dt} = -k(\theta_2-\theta_1) , \quad I_2 \frac{d\theta_2}{dt} = H_2 ,$$

where H is the angular momentum.

(b) Show that the characteristic equation is

$$\lambda^4 + \lambda^2 \left(\frac{k}{I_1} + \frac{k}{I_2}\right) = 0 ,$$

and find the solution corresponding to the non-zero eigenvalues. Interpret physically the mode corresponding with the non-zero eigenvalues [Hint: under no external torque the angular momentum is a constant, possibly zero].

(c) Show that

$$\begin{pmatrix}\theta_1\\ \theta_2\\ H_1\\ H_2\end{pmatrix} = \begin{pmatrix}\theta_0\\ \theta_0\\ 0\\ 0\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\theta_1\\ \theta_2\\ H_1\\ H_2\end{pmatrix} = \begin{pmatrix}\omega_0 t\\ \omega_0 t\\ \omega_0 I_1\\ \omega_0 I_2\end{pmatrix}$$

are also solutions. What are their physical interpretations?

2.7 For the network shown in Figure 2.7 obtain the system of differential equations for the currents through the resistors having resistances $\frac{R}{4}$, $\frac{3R}{2}$. Find the two relaxation configurations and the general solution.

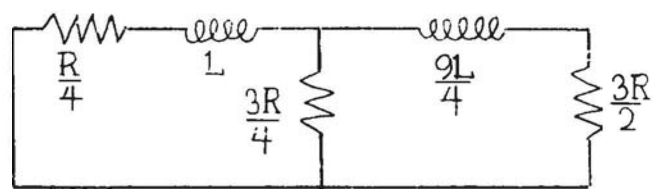


Figure 2.7

2.8 Three tanks are arranged so that when water is fed into the first tank an equal quantity of solution overflows from the first to the second tank, the second to the third, and from the third tank out of the system. Stirring devices keep the concentration in each tank uniform. The tanks have volume V_1 = 10,000 gal., V_2 = 8,000 gal. and V_3 = 5,000 gal. At the start, each tank is full of a solution having concentration C_0 lb./gal. Water is run into the first tank at the rate of 50 gal./min. and overflows as described above. See Figure 2.8.

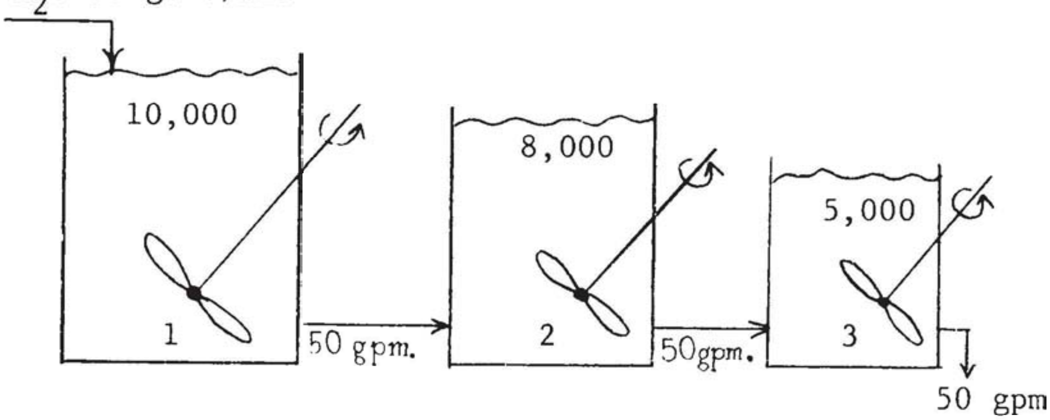


Figure 2.8

If C_1 is the concentration in lb./gal. of material in tank 1 at any time t, then 10,000 C_1 is the lb. of material in tank 1 at any time t. No material enters tank 1 because only water flows in. However, material leaves tank 1 at a rate of 50 C_1 lb./min. Thus, a material balance on tank 1 at any time gives

$$\frac{d}{dt}(10{,}000\ C_1) = -\ 50\ C_1.$$

(a) Determine the material balance differential equations for tanks 2 and 3. Show that the eigenvalues are -1/200, -1/160, -1/100.

(b) Find the time required to reduce the concentration in the first tank to $C_0/10$ and find the concentrations in the other two tanks at this time.
Answer: 200 log 10 = 460 min, $C_2 = 0.274C_o$, $C_3 = 0.414C_o$.

2.9 (a) Show that the matrix

$$(2.53) \qquad A = \begin{pmatrix} p & 1 \\ 0 & p \end{pmatrix}$$

has only one eigenvalue, $\lambda = p$, and only one independent eigenvector $v = c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Hence the methods of this section cannot provide a complete solution of the system $\frac{dx}{dt} = Ax$.

(b) Solve the uncoupled system

$$\frac{dx_1}{dt} = px_1 + x_2\ , \qquad \frac{dx_2}{dt} = px_2\ ,$$

by the methods of Chapter 1, Section 6, to get the general solution

$$x_1 = c_1 te^{pt} + c_2 e^{pt} ,$$

$$x_2 = c_1 e^{pt} ,$$

or, in vector form,

$$x = c_1 \begin{pmatrix} t \\ 1 \end{pmatrix} e^{pt} + c_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{pt}.$$

2.10 The system of equations,

$$(2.14) \qquad \frac{dx}{dt} = Ax ,$$

can be put in other forms by changes of variable of the type

$$y = Cx \quad \text{or} \quad x = C^{-1}y ,$$

where C is a non-singular matrix of constants. Show that with this change of variables (2.14) becomes

$$\frac{dy}{dt} = By,$$

where $B = CAC^{-1}$ (cf. Chapter 5, Section 7).

2.11 (a) If A is a 2x2 matrix, not a diagonal matrix, with only one eigenvalue $\lambda = p$, show that A has the form

$$A = \begin{pmatrix} p+a & b \\ c & p-a \end{pmatrix} ,$$

where $a^2 + bc = 0$ and either $b \neq 0$ or $c \neq 0$.

(b) If A has the above form show that if we take the matrix C of Problem 2.10 to be

$$C = \begin{pmatrix} -c & a+1 \\ c & -a \end{pmatrix} \text{ if } c \neq 0 \text{ , } \quad C = \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix} \text{ if } c = 0,$$

then

$$B = \begin{pmatrix} p & 1 \\ 0 & p \end{pmatrix} .$$

(c) Show how the above results can be combined with Problem 2.9 to solve the system (1.1) when there is only one distinct eigenvalue. Use this method to solve

$$\frac{dx_1}{dt} = 3x_1 + 18x_2 \text{ ,}$$

$$\frac{dx_2}{dt} = -2x_1 - 9x_2 .$$

Answer: $$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \left[c_1 \begin{pmatrix} -6t - 7 \\ 2t + 2 \end{pmatrix} + c_2 \begin{pmatrix} -6 \\ 2 \end{pmatrix} \right] e^{-3t} .$$

[Note. This method can be extended to the general case of an nxn matrix. The general form analogous to (2.53) is called the Jordan Canonical Form of a matrix.]

2.12 Use the above methods to solve the remaining case of Problem 1.3, the so-called "critically damped" case when

$$\left(\frac{R}{2L}\right)^2 = \frac{1}{LC} .$$

2.13 (a) By Euler's Method (Section 2 of Chapter 1) tabulate approximate solutions of the linear system with variable coefficients,

$$\frac{dx}{dt} = (1-t)x + ty \ ,$$

(2.54)

$$\frac{dy}{dt} = tx + (1-t^2)y \ ,$$

for the values of t: 0,0.1, 0.2,...,1.0.
For one solution $x_1(t)$, $y_1(t)$ use the initial conditions $x_1(0) = 1$, $y_1(0) = 0$; for another solution use $x_2(0) = 0$, $y_2(0) = 1$. [Note. You may find it convenient to do this on the computer.]

(b) By the comment following Theorem 2.1 any solution of (2.54) has the form

$$x(t) = c_1x_1(t) + c_2x_2(t) \ ,$$

$$y(t) = c_1y_1(t) + c_2y_2(t).$$

Tabulate the solution with boundary conditions $y(0) = 1$, $y(1) = 1$.

3. Two First Order Nonhomogeneous Simultaneous Differential Equations

From the work of the previous sections we conclude that a certain class of physical or mathematical problems can be described by equations (2.13) together with initial conditions (2.15). In physical problems, external energy is introduced by an initial disturbance and the system then performs what is called the free response. In the inertia-stiffness model of Example 1.1 the energy was initially stored in the "stiffness element" and the free response was a decay or relaxation because energy was dissipated in the resistors.

This situation occurs in systems ranging from the atomic scale to those of very large proportions -- including the Earth itself. One could imagine a large building or bridge as a mass-stiffness system and the free response would ensue after a disburbance arising from an earthquake or atmospheric blast wave. However, there are many physical situations in which the primary interest centers on the behavior while a source of external excitation acts. We admit certain types of excitation by including functions $g_1(t)$ and $g_2(t)$ in equations (2.1) for a system of two equations:

$$
\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + g_1 \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + g_2 \ ,
\end{aligned}
\tag{3.1}
$$

together with initial conditions

$$x_1(t_0) = X_1 \ , \ x_2(t_0) = X_2. \tag{3.2}$$

Such a system of equations is called <u>nonhomogeneous</u>, and the problem is to determine the solution to the nonhomogeneous system of equations.

In order to solve (3.1) we use an approach based on our knowledge of the behavior of similar single first order differential equations. We can write (3.1) in the vector form

$$\frac{dx}{dt} = Ax + g \ , \tag{3.3}$$

where, as in Section 2, x is the column vector with components x_1 and x_2, A is the matrix of coefficients a_{ij}, and g the column vector having components g_1 and g_2. From the work of Chapter 1, we know that if (3.3) were a single differential equation we could obtain the solution in two parts. With g zero we would have a homogeneous differential equation whose solution, called the complementary solution, is known to be of exponential form; it contains an arbitrary constant. With g present in the equation we seek a particular solution of the given differential equation. The total solution is the sum of the complementary and particular solutions, and the arbitrary constant is determined by the initial conditions. We will show that the same approach can be used to obtain the solution for the system of differential equations represented by the vector equation (3.3).

Example 3.1. The Blast Loading of Tall Structures

Tall structures such as stacks, elevated tanks, or skyscrapers can be subjected to sudden loading from the pressure front of an oncoming blast wave. The simplest idealization of such structures is a rigid mass supported on columns which possess stiffness when laterally deflected. Figure 3.1(a) shows an elevated water storage tank with incoming blast wave and Figure 3.1(b) shows the tank with deflection u caused by the lateral loading of the blast wave.

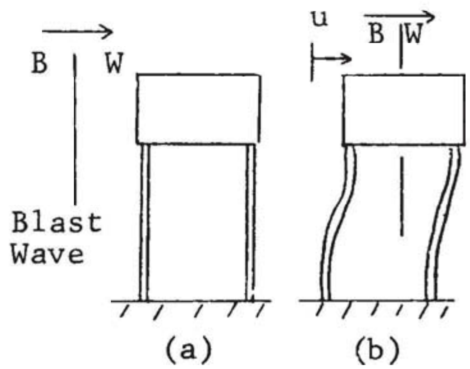


Figure 3.1

If the mass of the columns is small compared to that of the tank and its contents, then the columns are essentially springs whose stiffness would be determined by elementary beam theory. We will assume the supporting structure to have lateral stiffness k. Thus, we have the simple harmonic oscillator of Section 1, only now it is excited by a blast wave. The characterization of the blast loading has been the subject of extensive study and we will assume that, in the time interval of interest, the lateral force F is constant as shown in Figure 3.2.

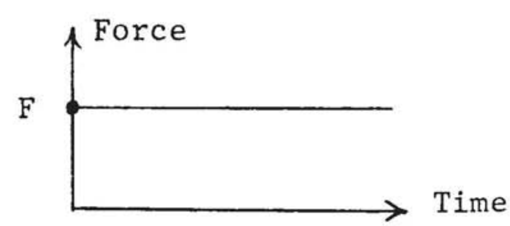


Figure 3.2

The mass m is subject to an exciting force F and a restoring force ku, hence Newton's law together with the definition of velocity give

$$(3.4)\qquad \begin{cases} m\,\dfrac{dv}{dt} = -\,ku + F\,, \\[2ex] \dfrac{du}{dt} = v\,, \end{cases} \qquad \text{or} \qquad \begin{cases} \dfrac{dv}{dt} = -\,\dfrac{k}{m}\,u + \dfrac{F}{m}\,, \\[2ex] \dfrac{du}{dt} = v\,. \end{cases}$$

Equations (3.4) are seen to be a special case of (3.1) with $x_1 = v$, $x_2 = u$, $a_{11} = 0$, $a_{12} = -\,k/m$, $g_1 = F/m$, $a_{21} = 1$ and $a_{22} = g_2 = 0$. As indicated earlier, the vector form of the equations suggests that we obtain the solution in two parts. We first solve the homogeneous set of equations for the complimentary solution $x_c = \begin{pmatrix} v_c \\ u_c \end{pmatrix}$. We obtain

$$(3.5)\qquad \begin{aligned} \frac{dv_c}{dt} &= -\,\frac{k}{m}\,u_c\,, \\ \frac{du_c}{dt} &= v_c\,, \end{aligned}$$

which are identical to (1.3) or (1.5). The solution, which can be written down immediately from (1.23), is

$$(3.6)\qquad \begin{aligned} v_c &= i\omega U_1 e^{i\omega t} - i\omega U_2 e^{-i\omega t}\,, \\ u_c &= U_1 e^{i\omega t} + U_2 e^{-i\omega t}, \end{aligned}$$

where U_1, U_2 are constants and $\omega = \sqrt{k/m}.$

A particular solution must now be found to the equations (3.4). Inspection of (3.4) suggests that because $g_1(y) = F$, a constant, there may be a solution of the form

(3.7) $$v_p = c_1 \ , \ u_p = c_2 \ ,$$

where c_1 and c_2 are constants. When substituted in (3.4) these values of v and u give

(3.8) $$\begin{aligned} 0 &= -\frac{k}{m} c_2 + \frac{F}{m} \ , \\ 0 &= c_1 \ , \end{aligned}$$

so that there is a particular solution

(3.9) $$v_p = 0 \ , \ u_p = \frac{F}{k} \ ,$$

of (3.4). We observe that the particular solution (3.9) is the static deflection of the spring, $\delta = F/k$, which would result from a very slowly applied force having final magnitude F.

If we add the complimentary solution (3.6) to the particular solution (3.9) we obtain

(3.10) $$\begin{aligned} v &= i\omega U_1 e^{i\omega t} - i\omega U_2 e^{-i\omega t} \ , \\ u &= \frac{F}{k} + U_1 e^{i\omega t} + U_2 e^{-i\omega t} \ , \end{aligned}$$

which contains constants U_1, U_2 and satisfies the original equations (3.4). In this example, the initial conditions are

(3.11) $$v(0) = 0 \ , \ u(0) = 0,$$

because the tank can be assumed to be at rest before arrival of the wave. When used in (3.10) these initial conditions give

$$0 = i\omega U_1 - i\omega U_2 \, ,$$

(3.12)

$$0 = \frac{F}{k} + U_1 + U_2 \, ,$$

which yield $U_1 = U_2 = -\frac{F}{2k}$. Consequently, solution (3.10) is

$$v = \frac{\omega F}{k} \; \frac{e^{i\omega t} - e^{-i\omega t}}{2i} \quad ,$$

(3.13)

$$u = \frac{F}{k} - \frac{F}{k} \; \frac{e^{i\omega t} + e^{-i\omega t}}{2} \quad ,$$

which in terms of trigonometric functions becomes

$$v = \frac{\omega F}{k} \sin \omega t \; ,$$

(3.14)

$$u = \frac{F}{k} (1 - \cos \omega t).$$

The response u is seen to be an oscillation, having amplitude F/k and natural frequency $f = \omega/2\pi = \sqrt{k/m}\,/2\pi$, which takes place about the static deflection F/k. Figure 3.3 shows this behavior. From an engineering viewpoint, the most important thing to observe is that the maximum displacement, 2F/k, is twice the static deflection F/k.

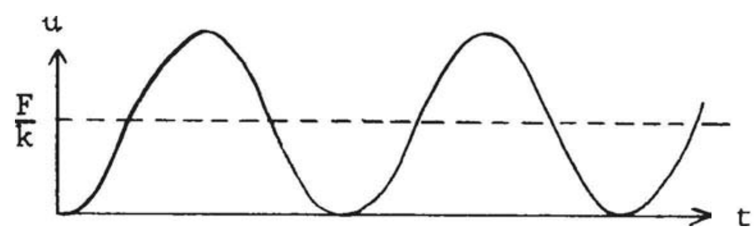


Figure 3.3

This result illustrates the inherent danger of shock loads, for design is frequently based on static loading. Relatively large deflections imply large strains which may cause failure.

The method of solution may be summarized as follows. The complementary solution is found and the form of a particular solution obtained by inspection. The total solution is then formed by superposition and <u>then</u> the initial conditions are applied in order to determine the constants of integration.

<u>Example 3.2</u>. Resonance

The phenomenon of resonance is inherent in any "inertia-stiffness" system to which is applied a periodic source of excitation. Because such systems oscillate naturally when disturbed, if an oscillatory external excitation is added, the possibility of reinforcement exists. Figures 3.4(a) and (b) show electrical and mechanical simple harmonic oscillators with sinusoidal voltage $V = V_0 \sin pt$ and force $F = F_0 \sin pt$ applied. The amplitudes V_0, F_0 are constant and p is the circular frequency of the source of excitation. We will use the language of the mechanical system, Figure 3.4(b), although by the analogy established in Section 1

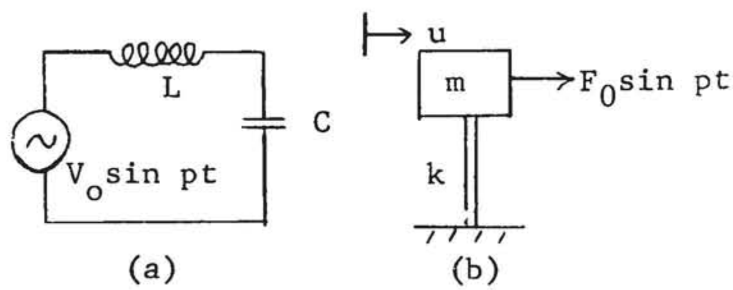


Figure 3.4

the analysis is applicable to either case.

Acting on the mass m is the exciting force $F_0 \sin pt$ and the restoring force $-ku$; hence the equations of motion become

$$(3.15) \qquad m\frac{dv}{dt} = -ku + F_0 \sin pt, \qquad \frac{du}{dt} = v.$$

We will assume the system to be at rest when the force is applied so that the initial conditions are those of (3.11). we again seek a solution to (3.15) with initial conditions (3.11) by finding the two parts of the solution - the complimentary and particular solutions. If we omit the term $F_0 \sin pt$ then the free response (complimentary solution) is of course identical with that previously obtained in Example 3.1, equations (3.6).

In seeking a particular solution we make use of a behavior which is characteristic of all physical systems governed by linear differential equations, i.e., those for which the physical model has been "linearized." Assume that the response due to the periodic excitation (particular solution) is itself periodic and of the same frequency as the excitation. Thus, take the particular integral in the form

$$(3.16) \qquad v_p = c_1 \sin pt + c_2 \cos pt, \qquad u_p = d_1 \sin pt + d_2 \cos pt,$$

where c_1, c_2, d_1 and d_2 are as yet undetermined constant coefficients. Observe that the circular frequency p in (3.16) is the same as that of the forcing function $F_0 \sin pt$. The coefficients c_1, c_2, d_1 and d_2 are determined by substituting (3.16) into (3.15). We obtain

$$
(3.17) \quad
\begin{aligned}
mp(c_1 \cos pt - c_2 \sin pt) &= -k(d_1 \sin pt + d_2 \cos pt) \\
&\quad + F_0 \sin pt, \\
p(d_1 \cos pt - d_2 \sin pt) &= c_1 \sin pt + c_2 \cos pt.
\end{aligned}
$$

Equations (3.17) will be satisfied for all values of t if in each equation the coefficients of sin pt and cos pt are equated. This gives the following four algebraic equations for c_1, c_2, d_1, and d_2 ,

$$
(3.18) \quad
\left\{
\begin{aligned}
mpc_1 &= -kd_2 , \\
-mpc_2 &= -kd_1 + F_0 , \\
pd_1 &= c_2 , \\
-pd_2 &= c_1 ,
\end{aligned}
\right.
\qquad \text{or} \qquad
\left\{
\begin{array}{rrrrl}
mpc_1 & & & +kd_2 & = 0 , \\
 & -mpc_2 & +kd_1 & & = F_0 , \\
 & -\ c_2 & +pd_1 & & = 0 , \\
-c_1 & & & -pd_2 & = 0.
\end{array}
\right.
$$

Solving the fourth equation for c_1 and substituting in the first gives

$$(mp^2 - k)d_2 = 0 ,$$

and hence, either $mp^2 - k = 0$ or

$$(3.19) \qquad d_2 = 0, \quad \text{and also} \quad c_1 = 0.$$

The second and third equations can be solved similarly, or one can use Cramer's Rule to get

$$c_2 = \frac{\begin{vmatrix} F_0 & k \\ 0 & p \end{vmatrix}}{\begin{vmatrix} -mp & k \\ -1 & p \end{vmatrix}} = \frac{pF_0}{k-mp^2}, \tag{3.20}$$

and

$$d_1 = \frac{\begin{vmatrix} -mp & F_0 \\ -1 & 0 \end{vmatrix}}{\begin{vmatrix} -mp & k \\ -1 & p \end{vmatrix}} = \frac{F_0}{k-mp^2}, \tag{3.21}$$

provided $k-mp^2 \neq 0$. The particular solution (3.16) now becomes

$$\begin{aligned} v_p &= \frac{pF_0}{k-mp^2} \cos pt , \\ u_p &= \frac{F_0}{k-mp^2} \sin pt. \end{aligned} \tag{3.22}$$

The fact that $c_1 = d_2 = 0$ in the assumed form of equations (3.16) is not surprising, for inspection of the first of equations (3.15) reveals that if all terms are to be sinusoidal, and hence of the same form as the force $F_0 \sin pt$, then u and v must be multiples of sin pt and cos pt respectively. When u and v are of this form the second of equations (3.15) can also be satisfied. Thus, in this example we could have replaced the assumed particular solution, equations (3.16), by taking $v_p = c_2 \cos pt$, $u_p = d_1 \sin pt$ directly.

We now add the particular solution (3.22) to the complimentary solution, (3.6), to obtain the general solution

$$\begin{aligned} v &= i\omega U_1 e^{i\omega t} - i\omega U_2 e^{-i\omega t} + \frac{pF_0}{k-mp^2}\cos pt, \\ u &= U_1 e^{i\omega t} + U_2 e^{-i\omega t} + \frac{F_0}{k-mp^2}\sin pt. \end{aligned} \tag{3.23}$$

The constants U_1 and U_2 are evaluated from the initial conditions (3.11) as follows

$$\begin{aligned} 0 &= i\omega U_1 - i\omega U_2 + \frac{pF_0}{k-mp^2}, \\ 0 &= U_1 + U_2, \end{aligned} \tag{3.24}$$

from which it follows that

$$U_1 = - U_2 = - \frac{pF_0}{2i\omega}\frac{1}{k-mp^2}. \tag{3.25}$$

If these values for U_1 and U_2 are inserted in (3.23) and the exponential functions replaced by sines and cosines, the solution becomes, finally,

$$\begin{aligned} v &= \frac{F_0}{k-mp^2}(-p\cos\omega t + p\cos pt), \\ u &= \frac{F_0}{k-mp^2}(-\frac{p}{\omega}\sin\omega t + \sin pt). \end{aligned} \tag{3.26}$$

The total response, (3.26), is the sum of the free oscillation (complimentary solution) of circular frequence ω

and a forced oscillation (particular solution) which has the same circular frequency p as the exciting force. Thus, the response is somewhat complicated and in general not periodic. Notice that because $\omega^2 = k/m$ the amplitude of the forced (and also the free) oscillation terms depends on the ratio

$$\frac{F_0}{k-mp^2} = \frac{F_0/k}{1-\left(\frac{p}{\omega}\right)^2}. \tag{3.27}$$

This amplitude is sensitive to the relative size of p compared to ω. If (p/ω) is small the ratio is nearly equal to F_0/k, the static deflection of the spring along. If $p \to \omega$ the amplitude becomes very large and this conditions is called <u>resonance</u>. Physically it is present because when $p \to \omega$ the driving force reinforces the natural (free) motion. Of course, in any real physical situation the amplitude cannot become infinite because as the amplitude gets large the "stiffness" element no longer retains its assumed linear characteristic. In some situations resonance must be avoided; however, if one wished a system to have a large response at a given frequency then resonance would be a way of achieving this. If $p = \omega$ the above method breaks down and a different particular solution must be found. Later we will determine the particular solution when $p = \omega$.

Figure 3.5, a plot of the amplitude given by (3.27), is called a frequency response diagram. Very often the absolute value of the amplitude

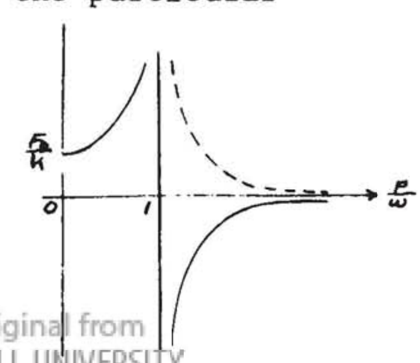


Figure 3.5

is plotted, so that for $p/\omega > 1$ the dashed curve is obtained. Notice that the sign of the amplitude is negative when $p/\omega > 1$; the particular response and the exciting force are in opposite directions and hence out of phase. For $(p/\omega) >> 1$ the amplitude drops toward zero because the system cannot follow the rapidly oscillating source of excitation.

The Particular Solution - Three Special Cases

Examples 3.1 and 3.2 are special cases of the system of two nonhomogeneous simultaneous equations (3.1). However, the forms of g(t), in these two examples, a constant or a sine function, occur frequently in applied science and we are led to consider the problem of determining a particular solution of (3.1) in specific cases. As indicated earlier, we use a method which consists of inspecting the form of g(t), using good judgment and experience in choosing a form for x_p, and then adjusting constants to achieve an identity when x_p is substituted into the differential equations (3.1).

Case I g(t) a constant. In this case equations (3.1) have the form

$$\frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + G_1 \,, \qquad (3.28)$$
$$\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + G_2 \,,$$

where G_1 and G_2 are given constants. We wish to select the form of $x_p = \begin{pmatrix} x_{1p} \\ x_{2p} \end{pmatrix}$ so that (3.28) are satisfied. We observe that if x_{1p} and x_{2p} are constant, then the terms involving the derivatives do not appear and we are left with a pair of linear equations which can be solved for x_{1p} and x_{2p}. Let

(3.29) $$x_{1p} = c_1 \quad , \quad x_{2p} = c_2 \ ,$$

where c_1 and c_2 are the unknown constants to be determined. Equations (3.28) can then be written as

(3.30) $$\begin{cases} a_{11}c_1 + a_{12}c_2 = -G_1 \\ a_{21}c_1 + a_{22}c_2 = -G_2 \end{cases} \quad \text{or} \quad Ac = -G.$$

Using Cramer's Rule we find c_1 and c_2 to be

(3.31) $$c_1 = \frac{\begin{vmatrix} -G_1 & a_{12} \\ -G_2 & a_{22} \end{vmatrix}}{|A|} \ , \ c_2 = \frac{\begin{vmatrix} a_{11} & -G_1 \\ a_{21} & -G_2 \end{vmatrix}}{|A|} \ ,$$

provided that the matrix A is nonsingular; i.e., det A $\neq$ 0. Thus, x_{1p} and x_{2p} given by (3.29) and (3.31) are the appropriate particular solutions when g_1 and g_2 are constant.

Example 3.3. The simple harmonic oscillator driven by the step force F of Example 3.1 has $a_{11} = 0$, $a_{12} = -k/m$, $a_{21} = 1$, $a_{22} = 0$, $G_1 = F/m$ and $G_2 = 0$. We note that det $A = k/m \neq 0$ and from (3.31) we conclude that $c_1 = 0$ and $c_2 = F/k$. This result coincides with that of (3.9) obtained previously.

Case II. g(t) of exponential form. We assume that $g_1(t)$ and $g_2(t)$ are of the same exponential form so that equations (3.1) become

$$\frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + G_1e^{pt} ,$$

(3.32)

$$\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + G_2e^{pt} ,$$

where G_1, G_2 and p are known constants. Observe that if x_{1p} and x_{2p} are also of the form ce^{pt} then every term of (3.32) has a factor e^{pt}, which then divides out of the equations, leaving a set of algebraic equations. Thus, assume

$$x_{1p} = c_1e^{pt} \quad , \quad x_{2p} = c_2e^{pt} , \tag{3.33}$$

where c_1 and c_2 are the undetermined coefficients to be determined. When (3.33) is substituted in (3.32) we obtain

$$pc_1e^{pt} = a_{11}c_1e^{pt} + a_{12}c_2e^{pt} + G_1e^{pt} ,$$

$$pc_2e^{pt} = a_{21}c_1e^{pt} + a_{22}c_2e^{pt} + G_2e^{pt} ,$$

from which e^{pt} can be cancelled to yield

$$\begin{cases} (p-a_{11})c_1-a_{12}c_2 = G_1, \\ -a_{21}c_1+(p-a_{22})c_2 = G_2, \end{cases} \quad \text{or} \quad (pI-A)c = G. \tag{3.34}$$

Again, Cramer's rule enables us to write

$$c_1 = \frac{\begin{vmatrix} G_1 & -a_{12} \\ G_2 & p-a_{22} \end{vmatrix}}{|pI-A|}, c_2 = \frac{\begin{vmatrix} p-a_{11} & G_1 \\ -a_{21} & G_2 \end{vmatrix}}{|pI-A|} , \tag{3.35}$$

provided that matrix (pI-A) is nonsingular. If det(pI-a) = 0 we could not obtain c_1 and c_2 and could not obtain a solution of the form (3.33). Observe that in finding the complementary solution when $G_1 = G_2 = 0$ we must find the roots of the characteristic equation $\det(\lambda I-A) = 0$. We conclude that we cannot obtain c_1 and c_2 if p coincides with one of the eigenvalues of A. Indeed in that case (3.33)is a solution of the homogeneous equation and does not represent a particular solution unless $G_1 = G_2 = 0$.

Example 3.4. In Example 3.1 the force due to the blast wave was idealized to the simplest form - a constant F shown in Figure 3.2. Field measurements show that blast forces follow very closely an exponentially decaying form Fe^{-pt}. The equations of motion are

$$\frac{dv}{dt} = -\frac{k}{m}u + \frac{F}{m}e^{-pt},$$

(3.36)

$$\frac{du}{dt} = v.$$

Comparison with (3.32) reveals that $a_{11} = 0$, $a_{12} = -k/m$, $a_{21} = 1$, $a_{22} = 0$, $G_1 = F/m$, $G_2 = 0$ and p is replaced by -p. Equations (3.35) give the coefficients c_1 and c_2 as

$$c_1 = \frac{-pF}{mp^2+k}, \quad c_2 = \frac{F}{mp^2+k}, \tag{3.37}$$

so that the particular solution of (3.36) is

$$v_p = \frac{-pF}{mp^2+k}e^{-pt},$$

(3.38)

$$u_p = \frac{F}{mp^2+k}e^{-pt}.$$

Case III. g(t) sinusoidal. First consider the case where $G_1(t)$ and $G_2(t)$ are complex functions $z_1(t)$ and $z_2(t)$ with real parts $R(z_1)$, $R(z_2)$ and imaginary parts $I(z_1)$, $I(z_2)$. Equations (3.1) become

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + R(z_1) + iI(z_1) \ , \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + R(z_2) + iI(z_2). \end{aligned} \tag{3.39}$$

We prove a very useful and general result by the following line of reasoning. We can expect that, in the particular solution, x_{1p} and x_{2p} will be complex functions so that

$$\begin{aligned} x_{1p} &= R(x_1) + iI(x_1), \\ x_{2p} &= R(x_2) + iI(x_2) \end{aligned} \tag{3.40}$$

If we substitute (3.40) into (3.39) we get

$$\left.\begin{aligned} &\frac{d}{dt}R(x_1) \\ &+i\ \frac{d}{dt}I(x_1) \end{aligned}\right\} = \left\{\begin{aligned} &a_{11}R(x_1)+a_{12}R(x_2)+R(z_1) \\ &+i[a_{11}I(x_1)+a_{12}I(x_2)+I(z_1)] \end{aligned}\right\} , \tag{3.41}$$

and

$$\left.\begin{aligned} &\frac{d}{dt}R(x_2) \\ &+i\ \frac{d}{dt}I(x_2) \end{aligned}\right\} = \left\{\begin{aligned} &a_{21}R(x_1)+a_{22}R(x_2)+R(z_2) \\ &+i[a_{21}I(x_1)+a_{22}I(x_2)+I(z_2)] \end{aligned}\right\} . \tag{3.42}$$

Each side of (3.41) and (3.42) is a complex number, and when two complex numbers are equal their real and imaginary parts are respectively equal. Hence, from the first lines of (3.41) and (3.42) we conclude that $x_{1p} = R(x_1)$, $x_{2p} = R(x_2)$ form a particular solution when $g_1 = R(z_1)$ and $g_2 = R(z_2)$. Similarly, from the second lines of (3.41) and (3.42) we conclude that $x_{1p} = I(x_1)$, $x_{2p} = I(x_2)$ form a particular solution when $g_1 = I(z_1)$ and $g_2 = I(z_2)$. Thus, if g_1 and g_2 are complex numbers and we somehow determine a particular solution x_{1p}, x_{2p} then we know that $R(x_1)$, $R(x_2)$ is a particular solution for $g_1 = R(z_1)$, $g_2 = R(z_2)$ and $I(x_1)$, $I(x_2)$ is a particular solution for $g_1 = I(z_1)$, $g_2 = I(z_2)$. This result greatly simplifies the algebra in determining particular solutions when $g_1(t)$ and $g_2(t)$ can be expressed as complex numbers, for the complex numbers can be written in exponential (polar) form.

Let the exciting force have the form F sin pt. Equations (3.1) become

$$(3.43) \qquad \begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + F_1 \sin pt \ , \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + F_2 \sin pt. \end{aligned}$$

Rather than attempt to solve (3.43) directly, we alter the problem to one where the exciting force is replaced by a complex function of the form

$$(3.44) \qquad g_1 = F_1 e^{ipt} \ , \ g_2 = F_2 e^{ipt} \ ,$$

so that the equations we must solve are

$$\frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + F_1e^{ipt} ,$$

(3.45)

$$\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + F_2e^{ipt}.$$

Because functions $g_1(t)$, $g_2(t)$ are of exponential form we can use the previous results of Case II. We know that x_{1p} and x_{2p} have the form of (3.33) so that

(3.46) $$x_{1p} = c_1e^{ipt} , \; x_{2p} = c_2e^{ipt}.$$

From (3.35) the constants c_1 and c_2 are

(3.47) $$c_1 = \frac{\begin{vmatrix} F_1 & -a_{12} \\ F_2 & ip-a_{22} \end{vmatrix}}{|ipI-A|} , \quad c_2 = \frac{\begin{vmatrix} ip-a_{11} & F_1 \\ -a_{21} & F_2 \end{vmatrix}}{|ipI-A|}$$

Notice that c_1 and c_2 are now complex numbers which can be written in polar form as

(3.48) $$c_1 = r_1e^{i\theta_1} , \; c_2 = r_2e^{i\theta_2}$$

where r_1, r_2, θ_1 and θ_2 can be found from (3.47). Substitution of c_1 and c_2 into (3.46) gives

(3.49) $$x_{1p} = r_1e^{i(pt+\theta_1)} , \; x_{2p} = r_2e^{i(pt+\theta_2)} .$$

If we take the imaginary parts of x_{1p} and x_{2p} we get

(3.50) $x_{1p} = r_1 \sin(pt+\theta_1)$, $x_{2p} = r_2 \sin(pt+\theta_2)$

which is the particular solution of (3.43), for the forcing functions of (3.43) are the imaginary parts of those of (3.45).

The method described above when $g_1 = F_1 \sin pt$ and $g_2 = F_2 \sin pt$ applies equally well for the situation when $g_1 = F_1 \cos pt$ and $g_2 = F_2 \cos pt$. Now, the real parts of (3.45) give $g_1 = F_1 \cos pt$ and $g_2 = F_2 \cos pt$ and we obtain a particular solution by taking the real parts of solution (3.49),

(3.51) $x_{1p} = r_1 \cos(pt+\theta_1)$, $x_{2p} = r_2 \cos(pt+\theta_2)$.

We conclude that if $g_1(t)$ and $g_2(t)$ are either of the form $C \sin pt$ or $C \cos pt$ we can determine constants c_1 and c_2 such that a particular solution x_{1p}, x_{2p} can be found. The constants c_1 and c_2 are found from (3.47) and we see that they exist if the denominator of $\det(ipI-A) \neq 0$. In the solution to the homogeneous equations, obtained by replacing $g_1(t)$, $g_2(t)$ by 0, the characteristic equation is $\det(\lambda I-A) = 0$. Again we see that should p be such that ip is one of the eigenvalues of A then $\det(ipI-A) = 0$ and we cannot obtain a solution by the method previously established. Physically, this corresponds to the situation when the circular frequency p of the forcing function equals a natural frequency ω and we are at resonance.

Example 3.5. The motion of a simple harmonic oscillator excited by the force $F_0 \sin pt$ is described by equations (3.15) of Example 3.2.

$$\frac{dv}{dt} = -\frac{k}{m}u + F_0 \sin pt,$$

(3.15)

$$\frac{du}{dt} = v.$$

Although we know that a particular solution of the form (3.16) exists, we wish to find this solution by using the techniques of Case III above. First, replace the exciting force by the exponential function $F_0 e^{ipt}$ so that the equations become

$$\frac{dv}{dt} = -\frac{k}{m}u + \frac{F_0}{m}e^{ipt}.$$

(3.52)

$$\frac{du}{dt} = v.$$

Of course, in doing this we do not make the physical force "complex," but merely define a new mathematical problem from whose solution, by taking the imaginary part, we can extract the solution of the physical problem. Following the approach given previously, we assume a solution of the form (3.46)

(3.53) $$v_p = c_1 e^{ipt}, \quad u_p = c_2 e^{ipt},$$

which when substituted into (3.52) gives

$$ipc_1 + \frac{k}{m}c_2 = \frac{F_0}{m},$$

$$c_1 - ipc_2 = 0.$$

When solved simultaneously these equations give

(3.54) $$c_1 = ip\frac{F_0/k}{1-(\frac{p}{\omega})^2}, \quad c_2 = \frac{F_0/k}{1-(\frac{p}{\omega})^2},$$

which when put in polar form (that is, when i is replaced by $e^{i\pi/2}$) become

$$(3.55)\qquad c_1 = p\,\frac{F_0/k}{1-(\frac{p}{\omega})^2}\,e^{i\frac{\pi}{2}}\ ,\quad c_2 = \frac{F_0/k}{1-(\frac{p}{\omega})^2}\ .$$

Thus, a particular solution of (3.52) is

$$(3.56)\quad v_p = p\,\frac{F_0/k}{1-(\frac{p}{\omega})^2}\,e^{i(pt+\frac{\pi}{2})}\ ,\quad u_p = \frac{F_0/k}{1-(\frac{p}{\omega})^2}\,e^{ipt}.$$

To obtain a particular solution to the original equations of motion (3.15), we take the imaginary part of (3.56),

$$(3.57)\qquad v_p = p\,\frac{F_0/k}{1-(\frac{p}{\omega})^2}\sin(pt+\frac{\pi}{2})\ ,\quad u_p = \frac{F_0/k}{1-(\frac{p}{\omega})^2}\sin pt,$$

and this result coincides with that obtained previously in equations (3.22). This particular example is sufficiently simple that one can use with equal ease either the direct approach of Example 3.2 or the complex number approach given here. However, in more complicated cases where $a_{11} \neq 0$, $a_{22} \neq 0$ and $g_2(t) \neq 0$ the complex number method is greatly to be preferred.

Although the three cases where g(t) is constant, exponential, or sinusoidal do not exhaust the possibilities for g(t), they are extremely important cases from a physical point of view. The method used to find a particular solution depends on the evaluation of constants c_1 and c_2; hence, the method is called the method of undetermined coefficients. For several other cases of $g(t)$ the particular solution can be found by the method of undetermined coefficients.

Since $C = Ce^{0t}$, $\sin pt = \mathrm{Im}(e^{ipt})$, and $\cos pt = \mathrm{Re}(e^{ipt})$, the results of this section can be summarized as follows: Given the system

$$\frac{dx}{dt} = Ax + Ge^{pt}, \tag{3.58}$$

where G is a constant vector, we try to find a constant vector C so that Ce^{pt} is a particular solution of (3.58). Since

$$\frac{d}{dt}(Ce^{pt}) = pCe^{pt} = pICe^{pt}$$

this leads us to the equation

$$(pI-A)C = G,$$

or

$$C = (pI-A)^{-1}G$$

provided p is not an eigenvalue of the matrix A.

Also, if $x = u + iv$ satisfies the system

$$\frac{dx}{dt} = Ax + g + ih,$$

where, A,u,v,g and h are real, then u satisifes

$$\frac{du}{dt} = Au + g$$

and v satisfies

$$\frac{dv}{dt} = Av + h.$$

Problems

3.1 Given the system of equations (see Problem 1.6)

$$\frac{dx}{dt} = -3x + 2y,$$

$$\frac{dy}{dt} = -2y + g_2(t),$$

(a) Find the solution when $g_2(t) = 6$, given the initial conditions $x(0) = 1$, $y(0) = 1$.

Partial Answer: $x = 3e^{-3t}-4e^{-2t}+2$.

(b) Find the solution when $g_2(t) = 2e^{2it}$, given the initial condition $x(0) = 0$, $y(0) = 0$.

Partial Answer: $y = \frac{1}{1+i}\left\{e^{2it}-e^{-2t}\right\}$.

3.2 For the system of equations of Problem 3.1 find the solution when $g_2(t) = 2\cos 2t$, but do not apply initial conditions.

Partial Answer: $y = Y_2e^{-2t}+\frac{1}{\sqrt{2}}\cos(2t - \frac{\pi}{4})$.

3.3 Consider the system of equations (see Problem 1.4)

$$\frac{dx_1}{dt} = x_1-2x_2+3e^{pt} ,$$

$$\frac{dx_2}{dt} = x_1-2x_2+6e^{pt}.$$

(a) If $p = 2$ find the solution with initial conditions $x_1(0) = -1$, $x_2(0) = 2$.

Answer: $x_1 = -3+2e^{-t}$; $x_2 = -\frac{3}{2}+2e^{-t}+\frac{3}{2}e^{2t}$.

(b) If $p = 0$ can a solution be found by the methods used thus far?

3.4 For the system of equations (see Problem 2.4)

$$\frac{dx_1}{dt} = -2x_1+2x_2+2\sin pt ,$$

$$\frac{dx_2}{dt} = -x_1-4x_2-5\sin pt ,$$

(a) Find a particular solution when $p = 4$.

Partial Answer: $x_2 = -\frac{2\sqrt{2}}{3}\sin(4t-\frac{\pi}{4})$.

(b) For which values of p will the method used in part (a) not yield a solution?

3.5 Find the solution to the system of equations

$$\frac{dx}{dt} = 2x-y ,$$

$$\frac{dy}{dt} = -3x+9 ,$$

with initial conditions $x(0) = 0$, $\left(\frac{dx}{dt}\right)_0 = -4$.

Partial Answer: $y = \frac{7}{4}e^{3t} - \frac{15}{4}e^{-t}+6$.

3.6 Find a particular solution for the system of equations

$$\frac{dx}{dt} = 2x+y ,$$

$$\frac{dy}{dt} = -5x+4y + \cos t.$$

Answer: $x_p = \frac{\sqrt{5}}{30}\cos(t+\theta_1)$, $y_p = -\frac{1}{6}\cos t$, where $\tan\theta_1 = \frac{1}{2}$, and θ_1 is acute .

3.7 Consider the system of equations

$$\frac{dx_1}{dt} = x_1 - 3x_2 + 6 ,$$

$$\frac{dx_2}{dt} = -2x_1 + e^{-3t}.$$

(a) Determine a particular solution in two steps by first finding a solution when $g(t) = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$ and then finding a solution when $g(t) = \begin{pmatrix} 0 \\ e^{-3t} \end{pmatrix}$. Show that the sum of these two solutions is a solution of the given system.

(b) Find the total solution to which initial conditions specifying $x_1(0)$, $x_2(0)$ could be applied.

(c) Can you make a general statement on how to find a particular solution of (3.1) when more than one functional form g(t) appears? (The question is vaguely expressed, make your answer precise!)

3.8 Given the system of equations

$$\frac{dx}{dt} = x + 3y + 8t ,$$

$$\frac{dy}{dt} = x - y.$$

(a) Find the form of a particular solution and evaluate any constants which appear in it.

(b) Find the solution which satisfies the initial conditions $x(0) = -2$, $y(0) = 4$.

3.9 Consider the general system of equations where $g(t) = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} t$, so that

$$\frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + G_1 t \ ,$$

$$\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + G_2 t \ ,$$

where G_1 and G_2 are constants. Determine the form of a particular solution and find expressions for any constants which appear in this solution

Partial Answer: Two of the constants are

$$\frac{\begin{vmatrix} -G_1 & a_{12} \\ -G_2 & a_{22} \end{vmatrix}}{|A|} \ , \quad \frac{\begin{vmatrix} a_{11} & -G_1 \\ a_{21} & -G_2 \end{vmatrix}}{|A|} \ .$$

3.10 Assume that a particular solution x_{1p}, x_{2p} is known for the system of equations

$$\frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + g_1(t) \ ,$$

$$\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 .$$

If the function $g_1(t)$ is multiplied by a number N, find the new particular solution in terms of the old. What do you conclude?

3.11 In Example 3.1, instead of a blast wave the structure is subject to an earthquake which causes a ground motion

$w = w_0 \sin pt$. See Figure 3.6. Find the system of equations analogous to (3.4), thus showing that the ground motion causes an exciting force $kw_0 \sin pt$ to act on the oscillator. Determine the solution and the amplitude of the forced oscillation. Is resonance possible in this case?

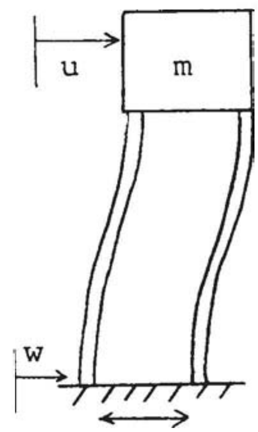


Figure 3.6

3.12 In Problem 1.3 the free response of a series L,C,R circuit was found when $(R/2L)^2 \neq 1/LC$. We wish to determine the response of this circuit to a constant voltage V_0 applied at time $t = 0$. Assume that initially the capacitor is uncharged and no current flows. Determine the system of equations for the current I and charge Q analogous to (3.4) and find their solution when $(R/2L)^2 < 1/LC$. For large values of t what values do I and Q approach and what are the voltage drops across each circuit element?

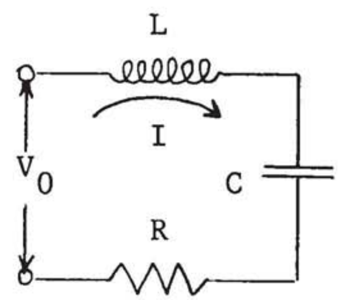


Figure 3.7

Partial Answer: $\begin{pmatrix} I \\ Q \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} \frac{\lambda_2 CV_0}{(\lambda_1 - \lambda_2)} e^{\lambda_1 t} - \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix} \frac{\lambda_1 CV_0}{(\lambda_1 - \lambda_2)} e^{\lambda_2 t} + \begin{pmatrix} 0 \\ CV_0 \end{pmatrix}.$

3.13 Consider the mass-spring-dashpot system shown in Figure 3.8. An ideal dashpot element exerts a resisting force cv, where c is the dashpot coefficient and v the velocity of relative motion between the plunger and enclosing cylinder. Derive the equations of motion for the velocity and displacement and show that they are analogous to those of Problem 3.12, only here the exciting force is cos pt rather than a constant.

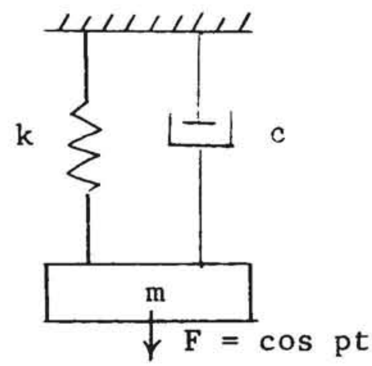


Figure 3.8

(a) Find a particular solution and thus determine the displacement amplitude and phase angle for the forced motion.

Partial Answer: Amplitude = $[(k-mp^2)^2 + (pc)^2]^{-1/2}$.

(b) Is resonance possible in the sense of (3.27)? Show resonance is possible in the sense that the amplitude of forced motion has a maximum provided $(k/m)>(c/\sqrt{2}m)^2$.

3.14 A simple harmonic oscillator having an inductance L in series with a capacitance C is acted upon by a voltage which increases linearly with time to a value V_0 at time T and then remains constant for t>T. The system is undisturbed before the voltage begins to act. (Figure 3.9.)

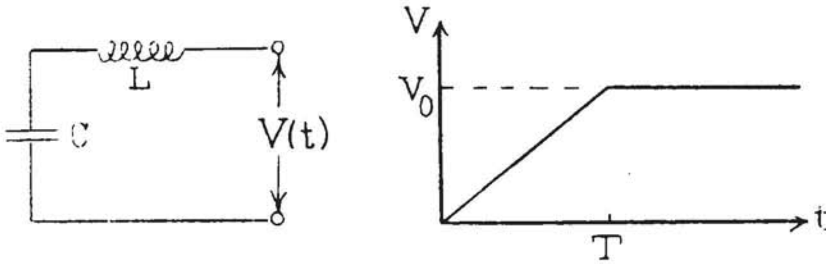


Figure 3.9

(a) For the time interval $0 \leqq t \leqq T$, derive the governing differential equations and determine the response. (See Problem 3.8 or 3.9.)

Answer: $$\begin{pmatrix} I \\ Q \end{pmatrix} = -\begin{pmatrix} iw \\ 1 \end{pmatrix} \frac{CV_o}{2iwT} e^{iwt} + \begin{pmatrix} -iw \\ 1 \end{pmatrix} \frac{CV_o}{2iwT} e^{-iwt}$$
$$+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{CV_0}{T} t + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{CV_0}{T}; \quad w = 1/\sqrt{LC} .$$

(b) When $t \geqq T$ derive the governing differential equations and state how you would find the initial conditions.

3.15 The resistor-inductor circuit shown in Figure 3.10 is initially undistrubed when a constant step voltage of 50 volts is applied across the terminals 1-2. Show that the equations governing currents i_1 and i_2 are

$$5\frac{di_1}{dt} = -30i_1+20i_2+50$$
$$5\frac{di_2}{dt} = 20i_1-30i_2.$$

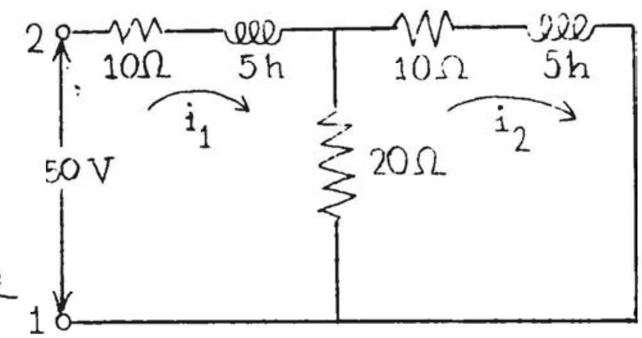


Figure 3.10

(a) For the complementary solution, determine the eigenvalues, corresponding eigenvectors, and "fast" and "slow" current configurations.

(b) Find a particular solution. Partial Answer: $i_{2p} = 2$.

(c) Find the solution which satisfies the initial conditions. What values do the currents approach for large t?

Partial Answer: $i_1 = -\frac{1}{2}e^{-10t} - \frac{5}{2}e^{-2t} + 3$.

(d) Determine the current in the 20Ω coupling resistor and sketch its time history.

4. Systems of First Order Nonhomogeneous Equations. General Case.

As in Section 2 where we generalized the system of two homogeneous equations of Section 1 to a system of n homogeneous equations, we now wish to generalize to n equations the work of Section 3 on two first order nonhomogeneous equations. Thus, we consider the system of equations

$$
\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + \ldots + a_{1n}x_n + g_1(t) \ , \\
\frac{dx_2}{dt} &= a_{21}x_1 + \ldots + a_{2n}x_n + g_2(t) \ , \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + \ldots + a_{nn}x_n + g_n(t) \ ,
\end{aligned} \tag{4.1}
$$

with initial conditions

$$x_1(t_0) = X_1 \ , \ x_2(t_0) = X_2 \ , \ldots, \ x_n(t_0) = X_n . \tag{4.2}$$

The system of equations (4.1) can be written in the form

$$\frac{dx}{dt} = Ax + g \ , \tag{4.3}$$

where $x(t)$ and $g(t)$ are to be regarded as functions whose values are n-component column vectors and A is the nxn matrix of coefficients a_{ij}. In this section we will discuss general properties of the solutions of (4.1) and determine the techniques to find these solutions for some important cases.

As in our previous considerations of the case n = 2, we construct the solution of (4.1) in two parts. First we consider the corresponding system of homogeneous equations

$$\frac{dx}{dt} = Ax \tag{4.4}$$

which were analyzed in Section 2. From Theorem 2.1 we know that the set of all solutions of (4.4) is an n-dimensional vector space and <u>any</u> solution can be written as a linear combination of a suitably chosen set of n vectors. Further, we find the n basic solutions by taking a solution in the form

$$x = ve^{\lambda t} \tag{4.5}$$

where v is a constant vector and λ a constant scalar. Substitution of (4.5) into (4.4) leads to the eigenvalue problem

$$\lambda v = Av \ ,$$

and if n independent eigenvectors of A can be found then n independent solutions of the form (4.5) are determined. If $x^1(t)$, $x^2(t), \ldots, x^n(t)$ are n independent solutions of (4.4) the general linear combination

$$x_c(t) = a_1 x^1(t) + a_2 x^2(t) + \ldots + a_n x^n(t),$$

where $a_1, a_2, \ldots, a_n$ are arbitrary constants, is called <u>the complementary solution of (4.3)</u>, for although it does not satisfy the given equation (4.3) it does satisfy the accompanying homogeneous equation (4.4). From Sections 2 and 3 we expect that the general solution of (4.3) is obtained by taking the sum of $x_c(t)$ and a particular solution $x_p(t)$ of (4.3). We now prove this.

<u>Theorem 4.1</u>. If $x_p(t)$ is a solution of (4.3) then any solution $x(t)$ of (4.3) is of the form

$$x(t) = x_p(t) + x_c(t) \tag{4.6}$$

where $x_c(t)$, the complementary solution, is a solution of (4.4).

<u>Proof</u>. We are given

$$\frac{dx}{dt} = Ax + g ,$$

$$\frac{dx_p}{dt} = Ax_p + g.$$

Subtracting gives

$$\frac{d(x-x_p)}{dt} = A(x-x_p) ,$$

and so $(x-x_p)$ is a solution x_c of (4.4); that is $x = x_p+x_c$.

This theorem indicates that the approach for solving (4.3) is the same as that used in the 2-dimensional case of Section 3. Thus, we first find the n independent solutions $x^1(t), \ldots, x^n(t)$ which make up x_c, the solution of the homogeneous equation, and then find one solution x_p of the nonhomogeneous equation (4.3). Theorem 4.1 says that any solution of (4.3) has the form

$$x = x_p + x_c = x_p + \sum_{i=1}^{n} a_i x^i \,, \tag{4.7}$$

where the constants $a_1, \ldots, a_n$ are determined so that x satisfies the initial conditions (4.2). We next consider some general theorems which are useful in determining x_p.

Theorem 4.2. The Principle of Superposition. Let the nonhomogeneous equations (4.1) have the form

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + \ldots + a_{1n}x_n + g_{11}(t) + g_{12}(t) + \ldots + g_{1m}(t) \,, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{dx_n}{dt} &= a_{n1}x_1 + \ldots + a_{nn}x_n + g_{n1}(t) + g_{n2}(t) + \ldots + g_{nm}(t) \,, \end{aligned} \tag{4.8}$$

so that in vector form (4.3) they become

$$\frac{dx}{dt} = Ax + \sum_{k=1}^{m} g_k . \tag{4.9}$$

Thus, g(t) is a sum of functions $g_k(t)$. Let x^k be a solution of

$$\frac{dx}{dt} = Ax + g_k .$$

Then $x(t) = \sum_{k=1}^{m} x^k(t)$ is a solution of (4.9).

Proof. We are given

$$\frac{dx^k}{dt} = Ax^k + g_k \,,$$

for k = 1, 2,...,m. Add these m equations to get

$$\frac{dx^1}{dt} + \frac{dx^2}{dt} + \dots + \frac{dx^m}{dt} = Ax^1 + Ax^2 + \dots + Ax^m + g^1 + g_2 + \dots + g_m \ ,$$

which can be written as

$$\frac{d}{dt}(x^1+x^2+\dots+x^m) = A(x^1+x^2+\dots+x^m) + \sum_{k=1}^{m} g_k$$

or

$$\frac{d}{dt}\sum_{k=1}^{m} x^k = A\sum_{k=1}^{m} x^k + \sum_{k=1}^{m} g_k \ .$$

Thus, if we take $x = \sum_{k=1}^{m} x^k$ it satisfies the equation

$$\frac{dx}{dt} = Ax + \sum_{k=1}^{m} g_k \ ,$$

which is equation (4.9).

Theorem 4.2 is useful in getting particular solutions when g(t), the exciting function, breaks naturally into the sum of two or more distinct parts. The theorem says we can find a particular solution x^k corresponding to the piece g_k and then obtain the total particular solution by adding the contributions x^k.

The next two theorems are helpful in handling the complex quantities that arise in connection with sines and cosines.

Theorem 4.3. Let A be a real matrix, let $g = g_R + ig_I$ and let $x = x_R + ix_I$ be a solution of (4.3); g_R, g_I, x_R and x_I are all real functions of t. Then x_R is a solution of $\frac{dx}{dt} = Ax+g_R$ and x_I is a solution of $\frac{dx}{dt} = Ax+g_I$.

<u>Proof</u>. We are given that

$$\frac{d(x_R+ix_I)}{dt} = A(x_R+ix_I)+g_R+ig_I ,$$

or

$$\frac{dx_R}{dt} + i\,\frac{dx_I}{dt} = Ax_R+iAx_I+g_R+ig_I.$$

Since A is real, equating real and imaginary parts gives

$$\frac{dx_R}{dt} = Ax_R+g_R ,$$

$$\frac{dx_I}{dt} = Ax_I+g_I ,$$

which completes the proof.

<u>Theorem 4.4</u>. If A is real and $x = x_R+ix_I$ is a solution of the homogeneous equation (4.4), x_R and x_I being real, then x_R and x_I are each solutions of (4.4).

<u>Proof</u>. This is just the special case of Theorem 4.3 in which $g = 0$.

<u>Corollary</u>. If $x_R + ix_I$ is a solution so is $x_R - ix_I$.

We are now ready to attack the solution of the system of equations (4.1) or (4.3). Theorem 4.1 shows that if we find the complementary solution, x_c, and a particular solution, x_p, then the sum of these two parts, (x_c+x_p), is the general solution of (4.3). The complementary solution is assumed known from the work of Section 2. The particular solution will be determined by generalizing to n dimensions the method of undetermined coefficients used in Section 3 for the 2-dimensional case. Theorems 4.2 and 4.3 will prove useful in constructing a particular solution.

Example 4.1. Consider the system of equations

$$\frac{dx_1}{dt} = 4x_1 - 2x_2 - 2 + 5e^{-t} ,$$

$$\frac{dx_2}{dt} = x_1 + x_2 + 4 + e^{-t} ,$$

with initial conditions $x_1(0) = 3$, $x_2(0) = 1$. When written in the form (4.3) we have

$$(4.10) \quad \frac{dx}{dt} = Ax+g \ , \text{ where } A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} , \ g = \begin{pmatrix} -2 \\ 4 \end{pmatrix} + \begin{pmatrix} 5 \\ 1 \end{pmatrix} e^{-t}.$$

The associated homogeneous equation,

$$\frac{dx}{dt} = Ax = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} x ,$$

is that of Example 2.1, and hence the complementary solution is given by (2.23) as

$$(4.11) \qquad x_c = \begin{pmatrix} 2 \\ 1 \end{pmatrix} v_1 e^{3t} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} v_2 e^{2t} ,$$

where v_1 and v_2 are constants. As Theorem 4.2 suggests we find a particular solution of (4.10) in two parts. First we find a particular solution of

$$(4.12) \qquad \frac{dx}{dt} = Ax+g_1 = Ax+ \begin{pmatrix} -2 \\ 4 \end{pmatrix} .$$

Inspection shows that with g_1 constant, a constant particular solution $x_p{}^1$ will satisfy (4.12). Thus, we assume

$$x_p{}^1 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} ,$$

and substitution in (4.12) gives the following system of equations for c_1 and c_2 ,

$$0 = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} -2 \\ 4 \end{pmatrix} \text{ or } \begin{cases} 0 = 4c_1 - 2c_2 - 2 , \\ 0 = c_1 + c_2 + 4. \end{cases}$$

When solved these yield $c_1 = -1$, $c_2 = -3$; thus, $x_{p1} = \begin{pmatrix} -1 \\ -3 \end{pmatrix}$.
Next we find a particular solution of

$$\frac{dx}{dt} = Ax+g_2 = Ax+ \begin{pmatrix} 5 \\ 1 \end{pmatrix} e^{-t}. \tag{4.13}$$

Observe that if we substitute a solution of the form

$$x_p{}^2 = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} e^{-t}$$

in (4.13) then each term will have e^{-t} as a factor and it can be divided out. We get

$$\begin{pmatrix} -d_1 \\ -d_2 \end{pmatrix} e^{-t} = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} e^{-t} + \begin{pmatrix} 5 \\ 1 \end{pmatrix} e^{-t} ,$$

which, after e^{-t} is eliminated, gives the following two equations for d_1, d_2,

$$-d_1 = 4d_1 - 2d_2 + 5 ,$$

$$-d_2 = d_1 + d_2 + 1.$$

These equations yield $d_1 = -1$, $d_2 = 0$; thus, $x_p{}^2 = \begin{pmatrix}-1\\0\end{pmatrix} e^{-t}$. The Superposition Principle, Theorem 4.2, enables us to obtain a particular solution of (4.10) by adding $x_p{}^1$ and $x_p{}^2$; thus

$$x_p = x_p{}^1 + x_p{}^2 = \begin{pmatrix}-1\\-3\end{pmatrix} + \begin{pmatrix}-1\\0\end{pmatrix} e^{-t}.$$

Theorem 4.1 enables us to find the general solution of (4.10) by adding x_c and x_p as in (4.7) to obtain

$$(4.14) \qquad x = \begin{pmatrix}2\\1\end{pmatrix} v_1 e^{3t} + \begin{pmatrix}1\\1\end{pmatrix} v_2 e^{2t} + \begin{pmatrix}-1\\-3\end{pmatrix} + \begin{pmatrix}-1\\0\end{pmatrix} e^{-t}.$$

Lastly, we determine v_1 and v_2 from the initial conditions $x_1(0) = 3$, $x_2(0) = 1$; hence, (4.14) for $x(0)$ gives

$$\begin{pmatrix}3\\1\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix} v_1 + \begin{pmatrix}1\\1\end{pmatrix} v_2 + \begin{pmatrix}-1\\-3\end{pmatrix} + \begin{pmatrix}-1\\0\end{pmatrix},$$

which when solved for v_1 and v_2 gives $v_1 = 1$ and $v_2 = 3$. Solution (4.14) of the given equation with the given initial conditions becomes

$$x = \begin{pmatrix}x_1\\x_2\end{pmatrix} \text{ where } \begin{cases} x_1 = 2e^{3t}+3e^{2t}-1-e^{-t} , \\ x_2 = e^{3t}+3e^{2t}-3. \end{cases}$$

<u>Example 4.2</u>. A Dilution Problem

An industrial chemical process generates colored waste water and dust which is to be discharged into a river. If the concentration of colored particles is too high, then the river color becomes objectionable. The color concentration is reduced by dilution in two tanks which are initially filled with clean water.

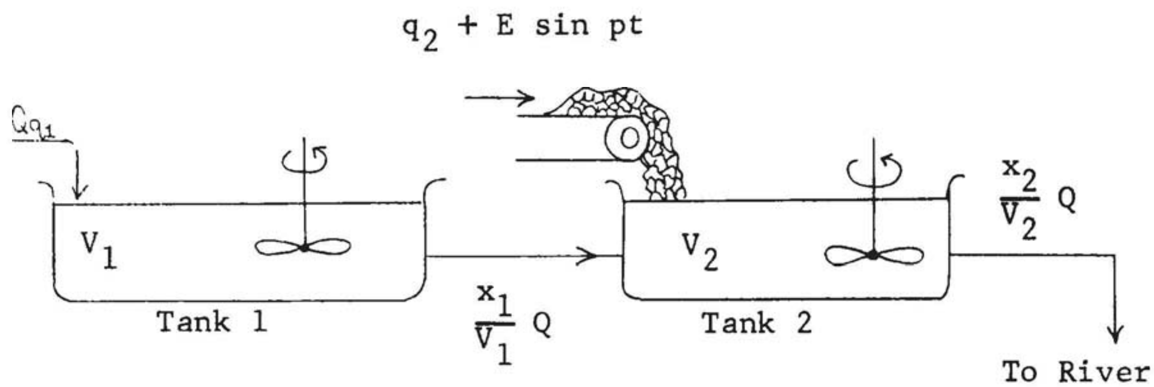


Figure 4.1

When the chemical process is put into operation waste water containing q_1 colored particles per gallon enters Tank 1 at a constant rate of Q gallons per minute. Tank 1 contains V_1 gallons in which perfect mixing is achieved by the stirring device. The contents of Tank 1 are pumped at a rate of Q gallons per minute into Tank 2 so that the tanks are coupled by this flow. Thus colored particles enter Tank 1 at a constant rate $q_1 Q$ particles per minute and leave at a rate $\frac{x_1}{V_1} Q$, where x_1 is the number of colored particles in Tank 1 at any time. Thus the rate of change of the number of colored particles in Tank 1 is

$$\frac{dx_1}{dt} = - \frac{Q}{V_1} x_1 + q_1 Q. \qquad (4.15)$$

In addition to the waste water fed into Tank 1, the chemical process includes waste colored dust (having negligible volume) which is fed into Tank 2 on a conveyor which pulsates so that it discharges (q_2+E sin pt) particles per minute. Tank 2 contains V_2 gallons and its contents are pumped into the river at a rate of Q gallons per minute. Thus colored particles leaving Tank 1 enter Tank 2 at an imput rate of $\frac{x_1}{V_1}$ Q and in addition the conveyor adds (q_2+E sin pt) particles per minute. Colored particles leave Tank 2 at the rate of $\frac{x_2}{V_2}$ Q, where x_2 is the number of colored particles in Tank 2 at any time. Thus the rate of change of the number of particles in Tank 2 is

$$\frac{dx_2}{dt} = - \frac{Q}{V_2} x_2 + \frac{Q}{V_1} x_1 + q_2 + E \sin pt. \tag{4.16}$$

The company wishes to determine the allowable amplitude E of dust pulsation rate so that the concentration of particles being discharged into the river never exceeds a level of S particles per gallon.

In order to determine the concentration of particles being discharged into the river, x_2/V_2, the system of equations (4.15), (4.16), must be solved. We have

$$\left\{\begin{aligned} \frac{dx_1}{dt} &= - \frac{Q}{V_1} x_1 \qquad\qquad + Qq_1 , \\ \frac{dx_2}{dt} &= \frac{Q}{V_1} x_1 - \frac{Q}{V_2} x_2 + q_2 + E \sin pt, \end{aligned}\right. \qquad \text{or } \frac{dx}{dt} = Ax+g , \tag{4.17}$$

with initial conditions

$$x_1(0) = 0 , \; x_2(0) = 0. \tag{4.18}$$

We obtain the complementary solution by solving the homogeneous equations

$$(4.19)\quad \begin{cases} \dfrac{dx_1}{dt} = -\dfrac{Q}{V_1} x_1 , \\ \dfrac{dx_2}{dt} = \dfrac{Q}{V_1} x_1 - \dfrac{Q}{V_2} x_2 , \end{cases} \qquad \text{or } \frac{dx}{dt} = Ax ,$$

with a solution of the form

$$(4.20)\quad \begin{cases} x_1 = X_1 e^{\lambda t} , \\ x_2 = X_2 e^{\lambda t} , \end{cases} \qquad \text{or } x = Xe^{\lambda t} ,$$

which when substituted in Equation (4.19) gives

$$(4.21)\quad \begin{aligned} (\lambda + \frac{Q}{V_1})\, X_1 \qquad\qquad &= 0 , \\ -\frac{Q}{V_1} X_1 + (\lambda + \frac{Q}{V_2})\, X_2 &= 0. \end{aligned}$$

The determinant of the coefficients gives the characteristic equation

$$(4.22)\quad (\lambda + \frac{Q}{V_1})(\lambda + \frac{Q}{V_2}) = 0 ,$$

so that the eigenvalues are $\lambda_1 = -Q/V_1$ and $\lambda_2 = -Q/V_2$. When $\lambda = \lambda_1$ equations (4.21) are satisfied if $X_2 = \dfrac{V_2}{V_1 - V_2} X_1$ so that the corresponding solution is

$$(4.23)\quad \begin{pmatrix} 1 \\ \dfrac{V_2}{V_1 - V_2} \end{pmatrix} e^{-\frac{Q}{V_1} t} .$$

When $\lambda = \lambda_2$ equations (4.21) are satisfied if $X_1 = 0$ so that the corresponding solution is

(4.24) $$\begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-\frac{Q}{V_2} t} .$$

Hence the complementary solution is a linear combination of (4.23) and (4.24),

(4.25) $$x_c = \begin{pmatrix} 1 \\ \frac{V_2}{V_1 - V_2} \end{pmatrix} c_1 e^{-\frac{Q}{V_1} t} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} c_2 e^{-\frac{Q}{V_2} t} .$$

We will determine the particular solution in three parts corresponding to the three inputs Qq_1, q_2 and E sin pt and use the Superposition Principle, Theorem 4.2, to obtain a particular solution to (4.17).(We separate Qq_1 and q_2 not for mathematical convenience but because we wish to determine their separate effects on the solution.) Thus we have

(4.26) $$g(t) = \begin{pmatrix} Qq_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ q_2 \end{pmatrix} + \begin{pmatrix} 0 \\ E \end{pmatrix} \sin pt = g_1 + g_2 + g_3 .$$

To find the solution of (4.17) with constant input g_1, substitute a particular solution of the form

(4.27) $$x_p{}^1 = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

and obtain the following simultaneous equations

$$
(4.28) \qquad
\begin{aligned}
0 &= -\frac{Q}{V_1} a_1 \qquad\qquad + Qq_1 , \\
0 &= \frac{Q}{V_1} a_1 - \frac{Q}{V_2} a_2 ,
\end{aligned}
$$

which have the solution $a_1 = V_1 q_1$ and $a_2 = V_2 q_1$. The particular solution (4.27) becomes

$$
(4.27) \qquad x_p^{\ 1} = \begin{pmatrix} V_1 q_1 \\ V_2 q_1 \end{pmatrix} .
$$

This result is easily interpreted, for with q_1 particles per gallon continuously entering Tank 1, eventually the steady state number of particles in Tank 1 will be $V_1 q_1$ and that in Tank 2, $V_2 q_1$.

To find a particular solution of (4.17) due to the constant rate of addition of dust particles from the conveyor, $g_2 = \begin{pmatrix} 0 \\ q_2 \end{pmatrix}$, assume a particular solution of the form

$$
(4.29) \qquad x_p^{\ 2} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} .
$$

Substitution of (4.29) gives the following simultaneous equations

$$
(4.30) \qquad
\begin{aligned}
0 &= -\frac{Q}{V_1} b_1 , \\
0 &= \frac{Q}{V_1} b_1 - \frac{Q}{V_2} b_2 + q_2 ,
\end{aligned}
$$

which have the solution $b_1 = 0$, $b_2 = \frac{V_2}{Q} q_2$. The particular solution (4.29) becomes

$$x_p{}^2 = \begin{pmatrix} 0 \\ \frac{V_2}{Q}\, q_2 \end{pmatrix}. \tag{4.31}$$

This result also can be interpreted, for the constant input q_2 does not reach Tank 1 and the steady-state rate of loss of particles from Tank 2, $\frac{x_2 Q}{V_2}$, must equal the rate of input q_2.

Lastly, we wish a particular solution of (4.17) which accounts for the fluctuating rate of input of dust particles, $g_3 = \begin{pmatrix} 0 \\ E \sin pt \end{pmatrix}$. In order to work with exponential functions replace g_3 by $\begin{pmatrix} 0 \\ Ee^{ipt} \end{pmatrix}$ so that equations (4.17) are replaced by

$$\begin{aligned} \frac{dx_1}{dt} &= -\frac{Q}{V_1}\, x_1 , \\ \frac{dx_2}{dt} &= \frac{Q}{V_1}\, x_1 - \frac{Q}{V_2}\, x_2 + Ee^{ipt}. \end{aligned} \tag{4.32}$$

All terms of (4.32) will involve e^{ipt} if we assume a particular solution

$$x_p{}^3 = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} e^{ipt}. \tag{4.33}$$

Substitution of (4.33) into (4.32) and division by e^{ipt} gives the equations

$$\begin{aligned} (ip + \frac{Q}{V_1})\, d_1 \qquad\qquad &= 0 , \\ -\frac{Q}{V_1}\, d_1 + (ip + \frac{Q}{V_2}) d_2 &= E , \end{aligned} \tag{4.34}$$

which clearly have the solution $d_1 = 0$, $d_2 = E/(ip + \frac{Q}{V_2})$, so that (4.33) becomes

$$(4.35) \qquad x_p{}^3 = \begin{pmatrix} 0 \\ E/(ip + \frac{Q}{V_2}) \end{pmatrix} e^{ipt}.$$

In order to extract the solution with the original $g_3 = \begin{pmatrix} 0 \\ E \sin pt \end{pmatrix}$ we use Theorem 4.3 and take the imaginary part of x_{p3} given by (4.35). We obtain

$$(4.36) \qquad x_p{}^3 = \begin{pmatrix} 0 \\ E\left[\left(\frac{Q}{V_2}\right)^2 + p^2\right]^{-1/2} \end{pmatrix} \sin(pt-\theta_3); \text{where } \tan\theta_3 = \left(\frac{p}{Q/V_2}\right).$$

As expected, solution (4.36) shows that Tank 1 is unaffected by the fluctuations in Tank 2, and the fluctuating number of particles in Tank 2 lags the input by phase angle θ_3.

Using the Superposition Principle, a particular solution of the original equations (4.17) is the sum of (4.27), (4.31) and (4.36)

$$(4.37) \qquad x_p = x_p{}^1 + x_p{}^2 + x_p{}^3.$$

By Theorem 4.1 the general solution of (4.17) is the sum of x_c and x_p; hence

$$x_1 = c_1 e^{-\frac{Q}{V_1}t} + V_1 q_1$$

$$x_2 = \frac{V_2}{V_1 - V_2} c_1 e^{-\frac{Q}{V_1}t} + c_2 e^{-\frac{Q}{V_2}t} + V_2 q_1 + \frac{V_2}{Q} q_2$$

$$+ \frac{E}{\sqrt{(\frac{Q}{V_2})^2 + p^2}} \sin(pt - \theta_3).$$

With no colored particles in either tank initially, initial conditions (4.18) when substituted in solution (4.38) give constants c_1 and c_2 as

$$(4.39) \qquad c_1 = -V_1 q_1 \ , \ c_2 = \frac{q_1 V_1 V_2}{V_1 - V_2} - V_2 q_1 - \frac{V_2}{Q} q_2 + \frac{E}{\sqrt{(\frac{Q}{V_2})^2 + p^2}} \sin \theta_3 .$$

Solution (4.38) shows that the number of colored particles in Tank 1 rises on an exponential curve from zero to the steady-state value $V_1 q_1$. Similarly, the number of colored particles in Tank 2 oscillates about a smooth rise from zero to the steady-state value $V_2 q_1 + \frac{V_2}{Q} q_2$. Solution (4.38) shows that the complementary solution is a transient which delays the attainment of the steady-state. The amplitude of oscillation depends not only on the amplitude E of the input, but also on the frequency, $f = p/2\pi$. Now, Q/V_2 is the frequency of "flushing" Tank 2 by circulating water at a rate Q gal./min. The amplitude of the oscillatory term can be written as

$$\frac{E}{\sqrt{(\frac{Q}{V_2})^2 + p^2}} = \frac{EV_2/Q}{\sqrt{1 + (\frac{p}{Q/V_2})^2}} ,$$

so that if $\frac{p}{Q/V_2} >> 1$ then the amplitude of the oscillatory response will be small; the input oscillation is so rapid that the tank cannot respond sufficiently fast. If $\frac{p}{Q/V_2} << 1$ then the amplitude of the oscillatory response would have its largest value EV_2/Q. Thus, the company should restrict E so that the maximum steady-state amplitude of x_2,

$$V_2 q_1 + \frac{V_2}{Q} q_2 + \frac{EV_2}{Q}.$$

does not exceed the allowable level.

Plots of the total response, equations (4.38), are shown in Figure 4.2.

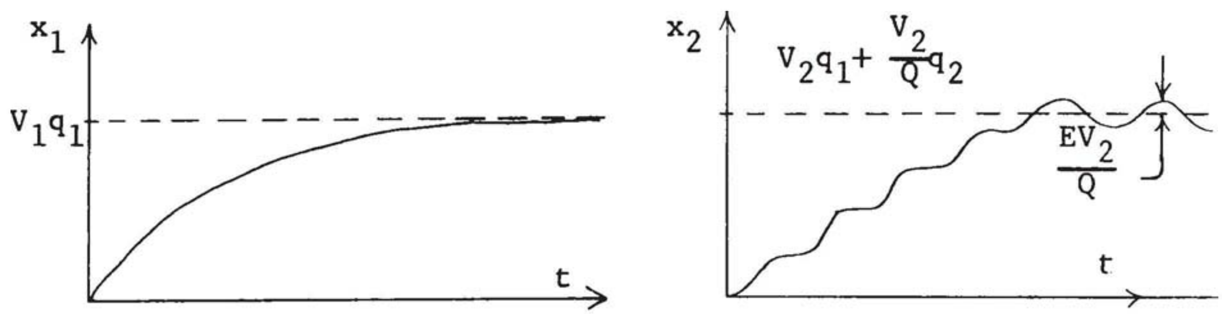


Figure 4.2

Problems

4.1 Find a particular solution to the system of equations,

$$\frac{dx_1}{dt} = -\ 2x_1 + 2x_2 + 1 + 17e^t\ ,$$

$$\frac{dx_2}{dt} = -\ x_1 - 4x_2 + 3 + 13e^{2t}\ ,$$

See Problem 3.4 for the complementary solution.

Answer:

$$x_p = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} + \begin{pmatrix} 5 \\ -1 \end{pmatrix} e^t + \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{2t}.$$

4.2 Find a particular solution to the system of equations,

$$\frac{dx_1}{dt} = x_1 - 4x_2 + x_3 + A \sin 2t \ ,$$

$$\frac{dx_2}{dt} = - 2x_2 + x_3 + B \cos 2t \ ,$$

$$\frac{dx_3}{dt} = + 4x_3 \ ,$$

where A and B are real.

Partial Answer:

$$x_p = \begin{pmatrix} -1/5 \\ 0 \\ 0 \end{pmatrix} A(\sin 2t + 2\cos 2t) + \begin{pmatrix} 3/5 \\ 1/4 \\ 0 \end{pmatrix} B \cos 2t + \begin{pmatrix} -1/5 \\ 1/4 \\ 0 \end{pmatrix} B \sin 2t .$$

4.3 Find a particular solution to the system of equations,

$$\frac{dx_1}{dt} = x_1 - 3x_2 + \sum_{m=1}^{p} (4m+3)e^{-4mt} \ ,$$

$$\frac{dx_2}{dt} = -2x_1 .$$

Partial Answer: $x = \sum_{m=1}^{p} \begin{pmatrix} 2m \\ 1 \end{pmatrix} \frac{e^{-4mt}}{(1-2m)}$.

4.4 Given the system of equations,

$$\frac{dx_1}{dt} + x_3 = 3 \cos 2t \ ,$$

$$\frac{dx_2}{dt} - x_3 = \cos 2t \ ,$$

$$\frac{dx_3}{dt} - x_1 = 0;$$

(a) Determine the complementary solution

Answer: $x_c = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} c_1 + \begin{pmatrix} 1 \\ -i \\ -i \end{pmatrix} c_2 e^{it} + \begin{pmatrix} 1 \\ -1 \\ i \end{pmatrix} c_3 e^{-it}.$

(b) Find a particular solution and check your result.

Answer: $x_p = \begin{pmatrix} 2 \sin 2t \\ 0 \\ - \cos 2t \end{pmatrix}$.

(c) If initial conditions $x_1(0) = 0$, $x_2(0) = 2$, $x_3(0) = 1$ must be satisfied, determine the solution

Answer: $x = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} -\cos t \\ \cos t \\ \sin t \end{pmatrix} + \begin{pmatrix} 2 \sin 2t \\ 0 \\ - \cos 2t \end{pmatrix}$.

4.5 Given the system of equations,

$$\frac{dx}{dt} = y + e^{pt} ,$$

$$\frac{dy}{dt} = z + 3 + e^{pc} ,$$

$$\frac{dz}{dt} = 13x - 17y + 5z + 2 + e^{pt};$$

(a) Find the complementary solution.

Answer:

$$x_c = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} c_1 e^t + \begin{pmatrix} 1 \\ 2+3i \\ -5+12i \end{pmatrix} c_2 e^{(2+3i)t} + \begin{pmatrix} 1 \\ 2-3i \\ -5-12i \end{pmatrix} e^{(2-3i)t}.$$

(b) Find a particular solution corresponding to $g_1(t) = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$.

(c) For what values of p can we <u>not</u> find a particular solution of exponential form? If p = 2 find a particular solution corresponding to $g_2(t) = \begin{pmatrix}1\\1\\1\end{pmatrix} e^{2t}$ and hence find a particular solution to the given equations.

Answer: $x_p = \begin{pmatrix}1\\0\\-3\end{pmatrix} + \begin{pmatrix}1\\1\\1\end{pmatrix} e^{2t}.$

(d) Determine the general solution of the given equations.

4.6 For the system of equations,

$$\frac{dx}{dt} = x + 3y + 2t \ ,$$

$$\frac{dy}{dt} = x - \ y + 2t \ ,$$

$$\frac{dz}{dt} = \quad - 6y + 4z \ ;$$

(a) Find a particular solution.

Answer: $\begin{pmatrix}-2t-1/2\\-1/2\\-3/4\end{pmatrix}.$

(b) With the complementary solution from Problem 2.3 determine the solution of the given equations which satisfies the initial conditions $x(0) = 2$, $y(0) = 1$, $z(0) = 3$.

Answer: $\begin{pmatrix}0\\0\\1\end{pmatrix} \frac{e^{4t}}{4} + \begin{pmatrix}1\\1/3\\1\end{pmatrix} 3e^{2t} - \begin{pmatrix}1\\-1\\-1\end{pmatrix} \frac{e^{-2t}}{2} + \begin{pmatrix}-2t-1/2\\-1/2\\-3/4\end{pmatrix}.$

4.7 For the R-L circuit with voltage $E(t)$ derive the appropriate differential equations.

(a) Find the complementary solution and describe

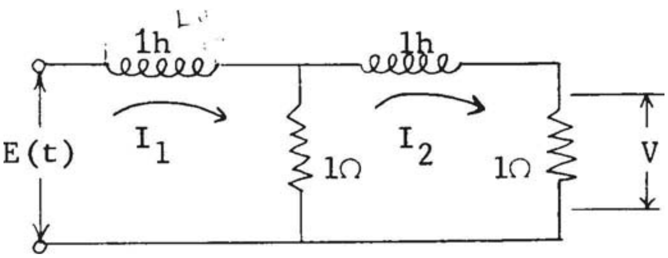


Figure 4.3

how it behaves for large t. What physical situation does the complementary solution describe?

(b) Determine the forced (steady state) voltage V if a constant voltage $E(t) = E_0$ is applied at time $t = 0$. Sketch the V(t) curve without actually solving for all constants.

(c) If a periodic voltage $E(t) = 3 \sin t$ is applied at time $t = 0$, determine the steady state voltage V.

4.8 A voltage $E = E_0 \sin pt$ is applied to the terminals of the resistor-inductor network of Example 2.3. Derive the governing differential equations in terms of the three loop currents I_1, I_2, and I_3. The complementary solution is given in Example 2.3. Does it contribute to the long time (steady state) response? Find the steady state response (particular solution) of the loop current I_3. Is resonance possible in this circuit? Briefly explain your answer on both a physical and mathematical basis. How does the amplitude of I_3 behave for high frequency excitation (p large)? Does I_3 lag or lead the applied voltage?

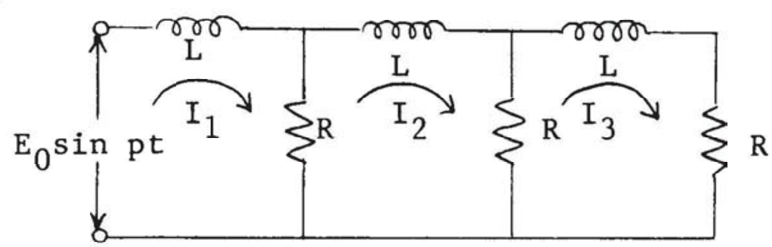


Figure 4.4

4.9 In the disk-shaft torsional system of Problem 2.6 assume that an external torque $T_0 \sin pt$ is applied to disk 1. Let the two moments of inertia be equal, $I_1 = I_2$. Determine the particular solution of the governing equations and plot the amplitude of disk 1 as a function of p. Is resonance possible?

4.10 Tanks A and B are connected as shown in Figure 4.5. When t = 0 tank A contains 90 gal. of brine having 50 lb. of salt in solution, while tank B has 60 gal. of water. Brine containing c lb./gal. of salt is pumped into tank A at the rate of $\frac{4}{3}$ gal./min.

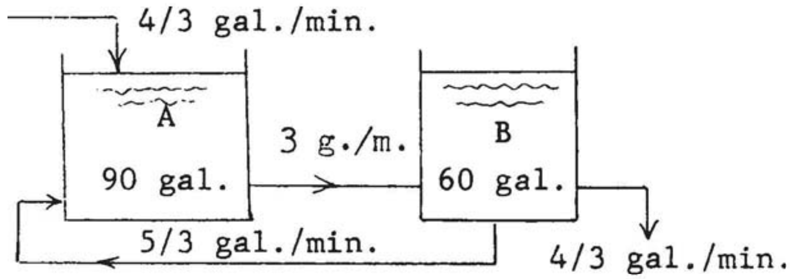


Figure 4.5

Mixture flows from A to B at a rate of 3 gal./min., from B to A at 5/3 gal./min., and out of B at 4/3 gal./min.

(a) Show that if x and y are the number of pounds of salt in tanks A and B respectively then the governing differential equations are

$$\frac{dx}{dt} = -\frac{1}{30}x + \frac{1}{36}y + \frac{4}{3}c\ , \quad \frac{dy}{dt} = \frac{1}{30}x - \frac{1}{20}y.$$

(b) Obtain the complementary solution. Interpret this solution physically after noting that if pure water is pumped into Tank A at the rate of $\frac{4}{3}$ gal./min. the homogeneous system of equations is unchanged.

Answer: $\begin{pmatrix} x \\ y \end{pmatrix} = C_1 \begin{pmatrix} 1 \\ -1.436 \end{pmatrix} e^{-.0732t} + C_2 \begin{pmatrix} 1 \\ .835 \end{pmatrix} e^{-.0101t}$.

(c) Find a particular solution.

(d) Show that there is a steady state condition (as $t \to \infty$). Could this have been predicted in advance?

4.11 Chemicals A and B must be mixed carefully because their reaction is highly exothermic. When brought together at an elevated temperature, α lb. of A combines instantaneously with 1 lb. of B to form $(1+\alpha)$ lbs. of a third substance C. Tank 1, containing volume V_1, is heated so that the reaction occurs in this tank. Tank 2, containing volume V_2, is kept cold so that no reaction can occur in it. Initially no B or C is present and Tank 1 contains m lbs. of A. A concentrated form of B (negligible volume) is bled into Tank 2 at the rate of b lb./min. The mixing process is accomplished by pumping r gal./min. of the mixture from Tank 2 into Tank 1, while r gal./min. of the contents of Tank 1 are pumped into Tank 2.

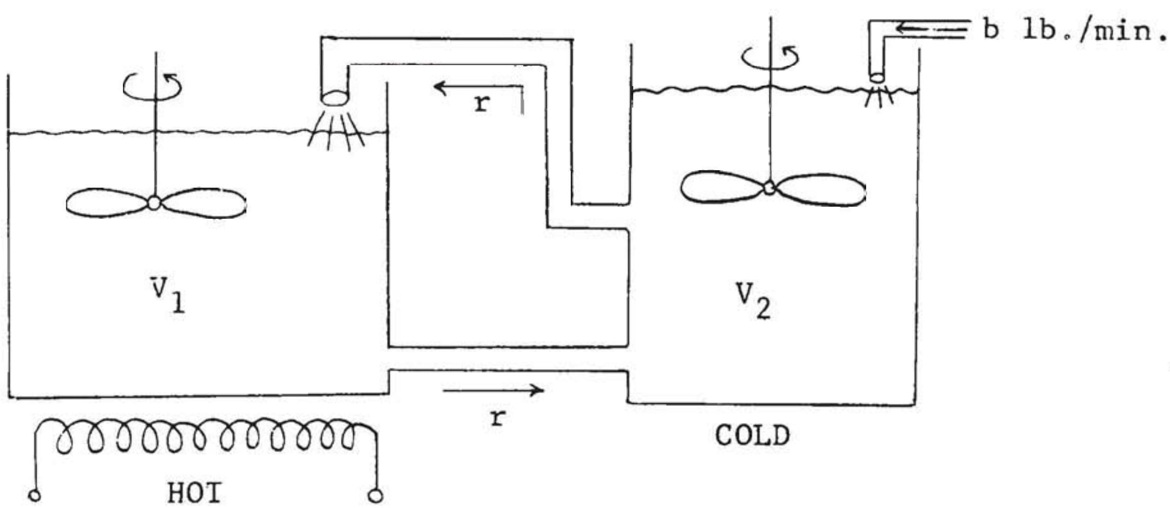


Figure 4.6

(a) Derive the equations for the rates of change of $A_1(t)$, $A_2(t)$, and $B_2(t)$ and find the complementary solution.

(b) Find a particular solution. [Hint. Try a solution of the form $x = a+bt$, where a and b are constant vectors.]

(c) Discuss the problem and its solution. Do the equations and the solution hold for arbitrarily large values of t?

4.12 Two masses M and m slide without friction on a horizontal plane. M is connected to a movable frame F by a spring that exerts a force k times the difference in the displacements of its two ends. M is connected to m by a shock absorber that exerts a force h times the difference in the velocities of its two ends; hence, h is the damping constant.

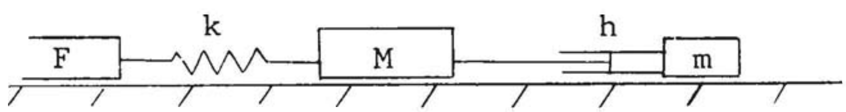


Figure 4.7

(a) Assuming the motion of F to be an oscillation of amplitude C and circular frequency p show that two of the equations of motion of the system are

$$M \frac{dv_1}{dt} = - hv_1 + hv_2 - kx_1 + kC \sin pt \ , \ \frac{dx_1}{dt} = v_1 \ ,$$

and derive the other two equations.

(b) What is the matrix A of the system, and the function g(t)? Show that the characteristic equation is

$$Mm\lambda^4 + h(M+m)\lambda^3 + km\lambda^2 + hk\lambda = 0.$$

Find the eigenvector associated with the zero eigenvalue and interpret it physically.

(c) If $m = 1$, $M = 2$, $k = \sqrt{6}$, $h = 2/\sqrt{6}$ show that $\lambda = -1$ is a second eigenvalue. Find the corresponding eigenvector.

(d) Are the two remaining eigenvalues real, imaginary, or complex? Hence, decide if resonance (in the sense of an infinite amplitude particular solution) is possible, without actually finding the particular solution.

5. Systems of Second Order Differential Equations

Although we have concentrated our attention thus far on systems of first order equations, frequently the need arises to consider systems of equations in which derivatives of order greater than one appear. We restrict our work to linear, constant coefficient equations of the form

$$\begin{aligned} \frac{d^2x_1}{dt^2} &= a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n+g_1(t) , \\ &\cdots\cdots\cdots\cdots \\ \frac{d^2x_n}{dt^2} &= a_{n1}x_1+a_{n2}x_2+\dots+a_{nn}x_n+g_n(t) , \end{aligned} \tag{5.1}$$

with initial conditions

$$\begin{aligned} x_1(t_0) &= X_1,\dots,x_n(t_0) = X_n , \\ \left(\frac{dx_1}{dt}\right)_{t_0} &= Y_1,\dots, \left(\frac{dx_n}{dt}\right)_{t_0} = Y_n . \end{aligned} \tag{5.2}$$

Although we will not use the approach here, it is clear that if in equations (5.1) and (5.2) we defined each first derivative as a new variable, $\frac{dx_i}{dt} = y_i$, then (5.1) and (5.2) would be a system of 2n first order linear equations in the 2n variables $x_1,\ldots,x_n$, $y_1,\ldots,y_n$. The theory developed previously in this chapter would apply directly and the solution could be found. However, it is better to make a fresh start.

Write (5.1) in the form

$$\frac{d^2x}{dt^2} = Ax + g \; , \tag{5.3}$$

and observe that an approach similar to that used previously appears promising. First suppress g and examine the homogeneous equation

$$\frac{d^2x}{dt^2} = Ax. \tag{5.4}$$

We will again call the general solution of (5.4) the complementary solution of (5.3) and attempt to get it as a linear combination of solutions of the form

$$x_c = ve^{\alpha t} \; , \tag{5.5}$$

where v is a vector of constants and α a scalar constant. We assume a solution of the form (5.5) because (just as in the case of a first order equation) we will find that we can divide out the factor $e^{\alpha t}$ after substituting from (5.5) into (5.4), and so replace the problem of solving a differential equation by a problem in linear algebra. In fact, on substituting we get

(5.6) $$\alpha^2 v e^{\alpha t} = A v e^{\alpha t} \quad \text{or} \quad \alpha^2 v = Av.$$

If in (5.6) we define $\alpha^2 = \lambda$ then (5.6) becomes

(5.7) $$\lambda v = Av$$

and we again are led to the eigenvalue problem for the matrix A. The characteristic equation

(5.8) $$\det(\lambda I - A) = 0$$

is an n^{th} degree polynomial in λ which we will assume has n distinct, non-zero roots $\lambda_1, \lambda_2, \ldots, \lambda_n$. Each such value of λ defines two values of α, namely $\alpha = \pm\sqrt{\lambda}$, so that we will assume that we have 2n distinct values of α, say $\alpha_1, \alpha_2, \ldots \alpha_{2n}$. Consequently, the homogeneous equation (5.4) has 2n solutions of the form (5.5).

From our previous discussion of the possibility of replacing (5.1) by a system of 2n first order linear equations it follows that every solution of the homogeneous system (5.4) is a linear combination of any set of 2n independent solutions. For the present we shall assume that the 2n solutions just obtained are independent. A proof of this fact, for an even less restrictive case than the one we are considering here, will be given later in this section. It follows then that any solution of (5.4) has the form

$$x_c = c_1 v_1 e^{\alpha_1 t} + \cdots c_{2n} v_{2n} e^{\alpha_{2n} t}, \tag{5.9}$$

where $v_1, \cdots, v_{2n}$ are eigenvectors and $c_1, \cdots, c_{2n}$ are scalars. Thus, (5.9) is the complementary solution of (5.3) and it contains the 2n constants $c_1, \ldots, c_{2n}$ which are determined by the 2n initial conditions (5.2) once the general solution of (5.1) is found.

Next we seek a particular solution, x_p, of (5.3) and by the Superposition Principle (Theorem 4.2 is easily extended) we construct the general solution of (5.3) by adding the complementary and particular solutions

$$x = x_c + x_p. \tag{5.10}$$

If initial conditions (5.2) are applied to the solution (5.10), the 2n constants c_i can be determined. The exact form of x_p in a given case depends, of course, on the form of g. As with the systems of first order equations previously studied, we will consider only cases where the form of x_p is evident by inspection and the method of undetermined coefficients can be used. Thus, the method of solution of (5.1) is very similar to that previously established for a system of first order equations.

Example 5.1. Deep Water Drilling - Mohole

When drilling in deep water it is necessary to connect lengths of drill pipe in order to reach the desired depths. The resulting long flexible rod is called the drill string and there are many dynamic problems associated with it. One of these is that the drill platform moves up and down due to wave action and excites the drill string into vertical oscillation. The drilling effectiveness is reduced and under some conditions the drill string breaks. Consider the case when two lengths of pipe are connected with a flexible coupling at C. Length 1 is flexibly mounted at the drive end A and the drill bit B on length 2 acts like a flexible connection to the ground. The exact stiffnesses at A, B, and C are not known, but we will assume those at A and B to be k while the coupling C has stiffness K. Thus, we are concerned with the vertical motion of two lengths, each having mass m, and the simplified model is shown in Figure 5.2.

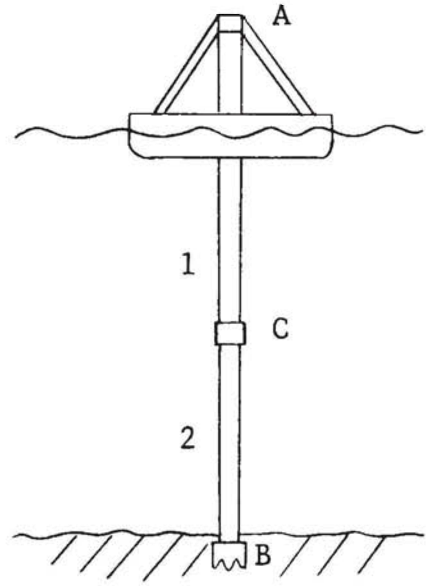


Figure 5.1

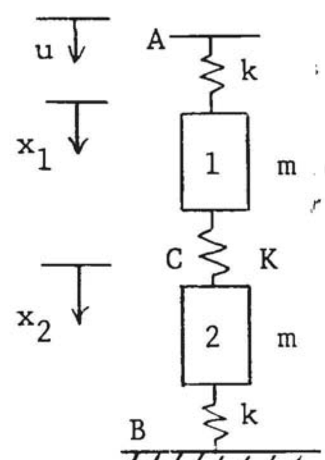


Figure 5.2

Define the following displacements:

u = downward platform deflection from equilibrium,

x_1= downward deflection of pipe 1 from equilibrium,

x_2= downward deflection of pipe 2 from equilibrium.

By considering the downward force due to the deflection across each spring, we can determine the force acting on each length of pipe. Because u, x_1, and x_2 are measured from their equilibrium positions, gravity does not appear in the calculation. Acting on length 1 is a downward force $k(u-x_1)$ from the spring at A and a downward force $K(x_2-x_1)$ from the spring at C. In terms of acceleration Newton's Law for length 1 gives

$$m \frac{d^2x_1}{dt^2} = k(u-x_1) + K(x_2-x_1). \tag{5.11}$$

Similarly, acting on length 2 is a downward force $K(x_1-x_2)$ from the spring at C and a downward force $-kx_2$ from the spring at B. Newton's Law for length 2 gives

$$m \frac{d^2x_2}{dt^2} = K(x_1-x_2) - kx_2. \tag{5.12}$$

If each equation is divided by m and the following definitions made,

$$\omega^2 = (k+K)/m \quad , \quad \eta^2 = K/m \ , \tag{5.13}$$

then (5.11) and (5.12) in form (5.1) become

$$\begin{aligned} \frac{d^2x_1}{dt^2} &= -\omega^2 x_1 + \eta^2 x_2 + \frac{k}{m} u \ , \\ \frac{d^2x_2}{dt^2} &= \eta^2 x_1 - \omega^2 x_2. \end{aligned} \tag{5.14}$$

Here we have

$$A = \begin{pmatrix} -\omega^2 & \eta^2 \\ \eta^2 & -\omega^2 \end{pmatrix} \ , \ g(t) = \begin{pmatrix} \frac{k}{m} \\ 0 \end{pmatrix} u(t).$$

Clearly, the platform displacement u(t), through its flexible connection at A, is the source of excitation.

To solve (5.14) we first find the complementary solution by solving the accompanying homogeneous equations. Let

$$x_c = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} e^{\alpha t} \ , \tag{5.15}$$

and substitute in the homogeneous equations to get

$$\begin{aligned} (\lambda+\omega^2)X_1 - \eta^2 X_2 &= 0 \ , \\ -\eta^2 X_1 + (\lambda+\omega^2)X_2 &= 0 \ , \end{aligned} \tag{5.16}$$

where we have put $\lambda=\alpha^2$. Equations (5.16) give the characteristic equation,

$$\begin{pmatrix} \lambda+\omega^2 & -\eta^2 \\ -\eta^2 & \lambda+\omega^2 \end{pmatrix} = (\lambda+\omega^2)^2 - \eta^4 = 0. \tag{5.17}$$

The two values of λ satisfying (5.17) are

$$\begin{aligned} \lambda_1 &= - \omega^2 + \eta^2 \ , \\ \lambda_2 &= - \omega^2 - \eta \ \ , \end{aligned} \tag{5.18}$$

and with the use of definitions (5.13) they become

$$\lambda_1 = -\ k/m\ ,$$

$$\lambda_2 = -\ (k+2K)/m.$$

If we make the additional definitions

(5.20) $$\omega_0{}^2 = k/m \quad , \quad s^2 = 1 + 2K/k\ ,$$

then we have

(5.21) $$\lambda_1 = -\ \omega_0{}^2\ ,\ \lambda_2 = -\ \omega_0{}^2 s^2.$$

The eigenvectors associated with each eigenvalue are found from (5.16). Using definitions (5.13) and $\lambda_1 = -k/m$ we get

$$\left[-\frac{k}{m} + \frac{(k+K)}{m}\right] X_1 - \frac{K}{m} X_2 = 0\ , \quad \text{or} \quad X_2 = X_1.$$

Thus, the eigenvalue λ_1 has associated with it the eigenvector

(5.22) $$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\ .$$

Similarly we find that the eigenvalue λ_2 has an associated eigenvector

(5.23) $$v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}\ .$$

We can now find the complementary solution. Because $\alpha = \pm\sqrt{\lambda}$ the two values of λ give rise to four values of α. Thus, from λ_1 in the form (5.21) we get

(5.24) $$\alpha_1 = +\omega_0 i \quad , \ \alpha_2 = -\omega_0 i.$$

Observe that the eigenvector (5.22) holds <u>for both values</u> α_1 and α_2, for equations (5.16) involve λ, and hence λ^2. Thus, we have two (independent) solutions of the form (5.15),

$$(5.25) \qquad \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{i\omega_0 t} \; , \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-i\omega_0 t} .$$

Similarly, from λ_2 the two values of α are

$$(5.26) \qquad \alpha_3 = +\omega_0 si \; , \; \alpha_4 = -\omega_0 si \; .$$

Again, the eigenvector (5.23) holds for both values α_3 and α_4 and we have two more (independent) solutions of the form (5.15),

$$(5.27) \qquad \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega_0 st} \; , \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-i\omega_0 st} .$$

A linear combination of the four solutions (5.25), (5.27) gives the general solution of the homogeneous equations; thus, the complementary solution is

$$(5.28) \quad x_c = \begin{pmatrix} 1 \\ 1 \end{pmatrix} c_1 e^{i\omega_0 t} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} c_2 e^{-i\omega_0 t} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} c_3 e^{i\omega_0 st} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} c_4 e^{-i\omega_0 st} .$$

Let us investigate the physical significance of the complementary solution. Equation (5.28) is a solution of the homogeneous equations; hence it is a solution

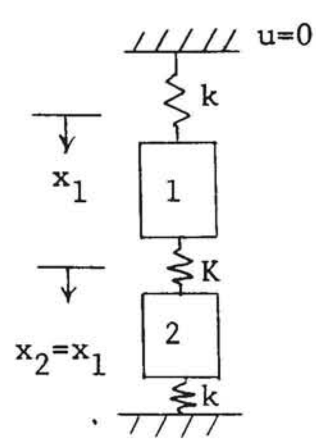


Figure 5.3

when $u = 0$ and there is no continuing external source of excitation. Thus, x_c is the free response of the system due to an initial disturbance only. First consider the result of an initial disturbance (see equations 5 2)

$$x_1(0) = x_2(0) = D_1 ,$$

$$\left(\frac{dx_1}{dt}\right)_0 = \left(\frac{dx_2}{dt}\right)_0 = 0.$$

When substituted in solution (5.28), these four conditions give $c_1 = c_2 = D_1/2$ and $c_3 = c_4 = 0$, so that (5.28) becomes

$$(5.29) \quad x_c = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{D_1}{2} (e^{i\omega_0 t} + e^{-i\omega_0 t}) = D_1 \begin{pmatrix} \cos \omega_0 t \\ \cos \omega_0 t \end{pmatrix}.$$

Clearly, $x_1 = x_2 = D_1 \cos \omega_0 t$, the two masses oscillate in unison with frequency $f = \frac{1}{2\pi}\sqrt{k/m}$, and the center spring of stiffness K is unstretched. This configuration (Figure 5.3) is just two uncoupled simple harmonic oscillators, each of mass m, stiffness k and natural frequency $f = \frac{1}{2\pi}\sqrt{k/m}$.

Similarly, if we give the system an initial disturbance

$$x_1(0) = - x_2(0) = D_2 ,$$

$$\left(\frac{dx_1}{dt}\right)_0 = \left(\frac{dx_2}{dt}\right)_0 = 0$$

then these four conditions give $c_1 = c_2 = 0$ and $c_3 = c_4 = D_2/2$ and solution (5.28) becomes

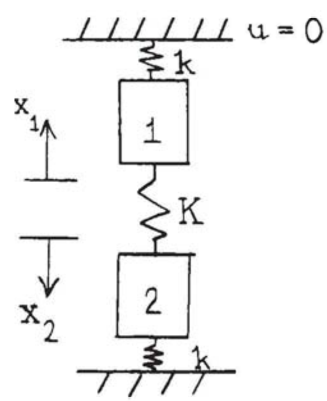


Figure 5.4

(5.30) $$x_c = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{D_2}{2} (e^{i\omega_0 st} + e^{-i\omega_0 st}) = D_2 \begin{pmatrix} \cos \omega_0 st \\ -\cos \omega_0 st \end{pmatrix}$$

In this case, $x_1 = -x_2 = D_2 \cos \omega_0 st$, the two masses oscillate in opposition to each other with frequency $f = \frac{1}{2\pi} \sqrt{\frac{(k+2K)}{m}}$, which is higher than the first frequency by the factor $s = \sqrt{1+ \frac{2K}{k}}$. The coupling spring K is now active and this configuration is shown in Figure 5.4.

The two configurations given by (5.29) and (5.30), and shown in Figures 5.3 and 5.4, are called normal modes of oscillation of the system, and the corresponding frequencies, the natural frequencies. Note that each eigenvector as an initial disturbance excites only its own normal modes.

Another form of the normal mode with circular frequency ω_0 can be obtained by starting with an initial velocity rather than an initial displacement, that is, with initial conditions

$$x_1(0) = x_2(0) = 0 ,$$

$$\left(\frac{dx_1}{dt}\right)_0 = \frac{dx_2}{dt} = V_1 .$$

This leads to the solution

(5.31) $$x_c = C_1 \begin{pmatrix} \sin \omega_0 t \\ \sin \omega_0 t \end{pmatrix}, \text{ where } C_1 = V_1/\omega_0 .$$

(Equation (5.31) can also be obtained from (5.29) by a shift in the time origin, i.e., by replacing t by $t - \frac{\pi}{2\omega_0}$.) The most general form of this normal mode is then a superposition of (5.29) and (5.31), namely

$$(5.32) \qquad x_c = \begin{pmatrix} 1 \\ 1 \end{pmatrix} (C_1 \sin \omega_o t + D_1 \cos \omega_0 t).$$

Similarly, the most general form of the second normal mode is

$$(5.33) \qquad x_c = \begin{pmatrix} 1 \\ -1 \end{pmatrix} (C_2 \sin \omega_0 st + D_2 \cos \omega_0 st).$$

Now by comparing the complementary solution (5.28) with the sum of the normal modes (5.32) and (5.33) it is easy to check that <u>any</u> complementary solution is expressible as a sum of normal modes of oscillation; we need only to set

$$D_1 = c_1 + c_2 \; , \; C_1 = - i(c_1 - c_2) \; ,$$

$$D_2 = c_3 + c_4 \; , \; C_2 = - i(c_3 - c_4).$$

Now that the complementary solution is known and its physical significance understood, we return to the problem of finding the solution to the governing equations (5.14). A particular solution cannot be determined until the platform motion $u(t)$ is given. We will assume that wind driven surface waves produce a steady oscillatory motion $u = U \sin pt$ of the platform. Thus,

$$(5.34) \qquad g(t) = \begin{pmatrix} \frac{k}{m} \\ 0 \end{pmatrix} U \sin pt \; ;$$

but we know from previous experience that the determination of a particular solution is facilitated by working with exponential functions. Therefore we replace (5.34) by

(5.35) $$g(t) = \begin{pmatrix} \frac{k}{m} \\ 0 \end{pmatrix} Ue^{ipt} ,$$

and take the imaginary part of the corresponding particular solution. To find a particular solution of (5.14) with g(t) given by (5.35), assume

(5.36) $$x_p = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} e^{ipt}.$$

Substitution in (5.14) gives the following two equations for b_1 and b_2,

(5.37) $$\begin{aligned} (-p^2+\omega^2)b_1 \qquad -\eta^2\, b_2 &= \frac{k}{m}\, U, \\ -\eta^2\, b_1 + (-p^2+\omega^2)b_2 &= 0. \end{aligned}$$

These equations have the solution

(5.38) $$b_1 = \frac{k}{m}\, U\, \frac{(\omega^2-p^2)}{\Delta} , \quad b_2 = \frac{k}{m}\, U\, \frac{\eta^2}{\Delta}$$

where Δ is the determinant of the coefficients on the left side of (5.37),

(5.39) $$\Delta = \begin{vmatrix} -p^2+\omega^2 & -\eta^2 \\ -\eta^2 & -p^2+\omega^2 \end{vmatrix} = (-p^2+\omega^2)^2-\eta^4.$$

In this case, b_1 and b_2 are real so that the imaginary part of (5.36) gives the particular solution as

(5.40) $$x_p = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \sin pt$$

where b_1 and b_2 are defined by (5.38).

It is important to observe that the platform oscillation U sin pt results in responses x_1 and x_2 of the same circular frequency p and amplitudes proportional to U. However, as with the simple harmonic oscillator, the amplitudes also depend on the frequency p and in particular on the denominator Δ. Comparison of equation (5.39) with the characteristic equation (5.17) reveals that Δ will be zero if $-p^2 = \lambda$. Hence from (5.21) we see that if p is close to ω_0 or $\omega_0 s$, either of the natural frequencies, the value of Δ will be small and an intolerably large forced oscillation will result. Thus the concept of resonance extends to the two-mass system and it occurs if the driving frequency coincides with either of the natural frequencies of the free response (complementary solution). Of course, if $p = \omega_0$ or $\omega_0 s$ then $\Delta = 0$ and a particular solution of the form (5.40) does not exist. To avoid resonance of the drill string, one would have to know the likely range of wave frequencies p, and then by design adjust the stiffnesses at A, B and C so that ω_0 and $\omega_0 s$ are not in this range. If this were not feasible, then energy dissipating devices would have to be included in the drill string.

We return now to the question of the linear independence of the 2n solutions of the form (5.5) that were used to get the complementary solution (5.9). Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of the matrix A, and $v_1, \ldots, v_n$ an associated set of eigenvectors. We shall assume that $v_1, \ldots, v_n$ are independent but we do not need to assume that $\lambda_1, \ldots, \lambda_n$ are distinct. Our conclusions will therefore apply to any symmetric matrix A, whether or not it has multiple eigenvalues. For the moment we do assume that no eigenvalue is zero, but we shall soon see how to remove this restriction also. Letting α_i and $-\alpha_i$ designate the two square roots of λ_i, we then have the 2n solutions

$$(5.41) \quad v_i e^{\alpha_i t}, \; v_i e^{-\alpha_i t}, \; i = 1, \ldots, n ,$$

and our problem is to show that these are independent.

By Theorem 6.1 of Chapter 2, the solutions (5.41) are independent if the only constants $c_1, \ldots, c_n, d_1, \ldots, d_n$ for which

$$(5.42) \qquad \sum_{i=1}^{n} c_i v_i e^{\alpha_i t} + \sum_{i=1}^{n} d_i v_i e^{-\alpha_i t} = 0$$

are $c_i = d_i = 0$ for $i = 1, \ldots, n$. If (5.42) is to hold for all values of t it must hold in particular for $t = 0$, and so

$$(5.43) \qquad \sum_{i=1}^{n} c_i v_i + \sum_{i=1}^{n} d_i v_i = \sum_{i=1}^{n} (c_i + d_i) v_i = 0.$$

We assumed that the v_i were independent; hence by the same Theorem 6.1 it follows from (5.43) that

$$(5.44) \qquad c_i + d_i = 0 , \; i = 1, \ldots, n.$$

Now since (5.42) is an identity in t we may differentiate both sides of the equation to obtain

(5.45) $$\sum_{i=1}^{n} \alpha_i c_i v_i e^{\alpha_i t} - \sum_{i=1}^{n} \alpha_i d_i v_i e^{-\alpha_i t} = 0.$$

Treating this equation as we treated (5.42) gives us

(5.46) $$\alpha_i (c_i - d_i) = 0 , \quad i = 1,\ldots,n.$$

Since $\alpha_i \neq 0$, $i = 1,\ldots,n$, we conclude from (5.44) and (5.46) that

(5.47) $$c_i = d_i = 0 , \ i = 1,\ldots,n.$$

Thus (5.41) holds only if (5.47) is true, and by Theorem 6.1 the solutions (5.41) are independent.

Finally, let us consider the case of an eigenvalue $\lambda = 0$. If v is an associated eigenvector then $Av = \lambda v = 0$. It is easy to check that

(5.48) $$x = v \quad \text{and} \quad x = vt$$

are solutions of (5.4), for in each case we see at once that $\frac{d^2x}{dt^2} = 0$ and $Ax = 0$. An argument similar to the one above can be given to prove that if some of the 2n solutions in (5.41) are of the type of (5.48) the total set is still independent. (See Problem 5.8)

Example 5.2. Find the general solution of

(5.49) $$\begin{aligned} \frac{d^2x}{dt^2} &= 5x - 2y + 4z , \\ \frac{d^2y}{dt^2} &= -2x + 8y + 2z , \\ \frac{d^2z}{dt^2} &= 4x + 2y + 5z. \end{aligned}$$

Here

$$A = \begin{pmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{pmatrix}$$

is a symmetric matrix, so we know that there are 3 independent eigenvectors. We find

$$\det(\lambda I - A) = \begin{vmatrix} \lambda-5 & 2 & -4 \\ 2 & \lambda-8 & -2 \\ -4 & -2 & \lambda-5 \end{vmatrix} = \lambda^3 - 18\lambda^2 + 81\lambda = \lambda(\lambda-9)^2 ,$$

and so the distinct eigenvalues are 0 and 9. For $\lambda = 0$ the equations

$$(\lambda I - A)v = 0, \text{ or } \begin{pmatrix} -5 & 2 & -4 \\ 2 & -8 & -2 \\ -4 & -2 & -5 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = 0 ,$$

have only one independent solution, which may be taken to be $\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$. For $\lambda = 9$ the corresponding equations reduce to the single equation

$$(5.50) \qquad 2v_1 + v_2 - 2v_3 = 0.$$

This has two independent solutions, for instance $\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

The general solution of (5.49) is then

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} (c_1+c_2t) + \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} (c_3e^{3t}+c_4e^{-3t}) + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} (c_5e^{3t}+c_6e^{-3t}).$$

Note that there is nothing special about our choice of $\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ as solutions of (5.50) - any two independent solutions will serve as well. In problems arising from physical situations there may be advantages in picking solutions that have simple physical interpretations (see Problem 5.7).

Problems

5.1 The free oscillations of a coupled two-mass system are described by the equations

$$\frac{d^2x}{dt^2} = - 3x + 2y ,$$

$$\frac{d^2y}{dt^2} = 6x - 7y.$$

(a) Determine the general solution of the equations. If x and y are the two displacements, identify the configurations in each of the normal modes and determine the two natural frequencies.
Partial Answer: $f_1 = \frac{3}{2\pi}$, $f_2 = \frac{1}{2\pi}$.

(b) If the system is initially excited by initial conditions

$$x(0) = 6,\ y(0) = -2,\ \left(\frac{dx}{dt}\right)_0 = 0,\ \left(\frac{dy}{dt}\right)_0 = 0 ,$$

evaluate the integration constants of part (a) and hence find the fraction of each normal mode required to meet these initial conditions.

Partial Answer: $x = 2 \cos 3t + 4 \cos t$.

5.2 Figure 5.5 shows two L-C loops coupled by capacitance K.

(a) Show that the equations governing free oscillations in the circuit are

$$L\frac{d^2Q_1}{dt^2} + \left(\frac{1}{C} + \frac{1}{K}\right)Q_1 - \frac{1}{K}Q_2 = 0 ,$$

$$L\frac{d^2Q_2}{dt^2} + \left(\frac{1}{C} + \frac{1}{K}\right)Q_2 - \frac{1}{K}Q_1 = 0.$$

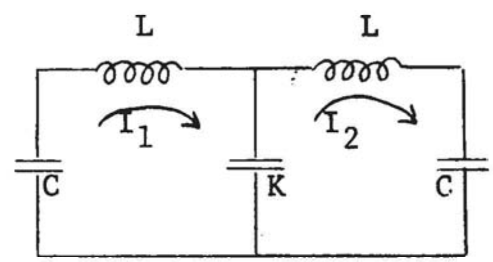


Figure 5.5

(b) Determine the natural frequencies of oscillation and the corresponding normal mode configurations.

(c) Draw a mechanical system which is the analogue of the electrical circuit.

5.3 (a) For the shaft-flywheel system of Problem 2.6 express the equations governing the motion in terms of two second order equations in θ_1 and θ_2. Determine the characteristic equation, the normal mode configuration and corresponding natural frequency for the non-zero eigenvalue.

(b) Determine the normal mode configuration and related solutions for the zero eigenvalue.

(c) Write the general solution of the equations.

(d) Find initial conditions which would result in motion involving just the normal mode of part (a). Check your answer.

5.4 Figure 5.6 shows an idealized three compartment space station interconnected by two light passageways which act as springs of stiffness k along the direction of connection. It is important to know the natural frequencies of vibration so that environmental sources of excitation, such as periodic internal moving parts, do not cause resonance which leads to fatigue failure. Center compartment 1 has mass M and the other two compartments have mass m. Assume the space station to be sufficiently remote from any external sources of excitation.

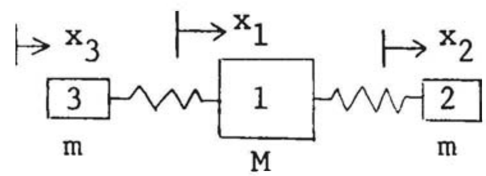


Figure 5.6

(a) With displacements x_1, x_2, x_3 measured from the equilibrium positions, draw a free body diagram of each mass and show that the governing equations of motion are

$$M \frac{d^2 x_1}{dt^2} = -2kx_1 + kx_2 + kx_3 ,$$

$$m \frac{d^2 x_2}{dt^2} = kx_1 - kx_2 ,$$

$$m \frac{d^2 x_3}{dt^2} = kx_1 - kx_3 .$$

(b) Show that the characteristic equation is

$$(\lambda + \frac{2k}{M}) (\lambda + \frac{k}{m})^2 - 2 \frac{k^2}{mM} (\lambda + \frac{k}{m}) = 0$$

and thus find the eigenvalues and corresponding eigenvectors.

Partial Answer: $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} -2m/M \\ 1 \\ 1 \end{pmatrix}$

(c) Determine the normal mode configurations and natural frequencies for the two non-zero eigenvalues. Prove that the conservation of linear momentum is satisfied in each normal mode.

(d) What are the physical interpretations of the two solutions arising from the zero eigenvalue?

(e) From the previous parts, construct the general solution.

(f) A space station proportioned such that M = 4m enters a meteroid shower, and compartment 3 is struck and given an instantaneous velocity V so that initial conditions (5.2) become

$$x_1(0) = x_2(0) = x_3(0) = 0, \text{ or } x(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} ,$$

$$\left(\frac{dx_1}{dt}\right)_0 = \left(\frac{dx_2}{dt}\right)_0 = 0 , \quad \left(\frac{dx_3}{dt}\right)_0 = V, \text{ or } \left(\frac{dx}{dt}\right)_0 = \begin{pmatrix} 0 \\ 0 \\ V \end{pmatrix} .$$

Write the six equations from which the six constants can be determined and show that the response of the space station to the meteoroid impact is

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \frac{V}{3} t - \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \frac{V}{2\sqrt{\frac{k}{m}}} \sin \sqrt{\frac{k}{m}}\, t + \begin{pmatrix} -.5 \\ 1 \\ -.5 \end{pmatrix} \frac{V}{3\sqrt{\frac{3k}{2m}}} \sin \sqrt{\frac{3k}{2m}}\, t \ .$$

5.5 Figure 5.7(a) shows three masses each of mass m attached equidistantly to a string which has a high initial tension T. The system can oscillate about its equilibrium position, for Figure 5.7(b) shows that tensions T have vertical components which act as springs attached to the masses. We assume that the displacements are relatively small so that T remains constant. From Figure 5.7(b) we can write

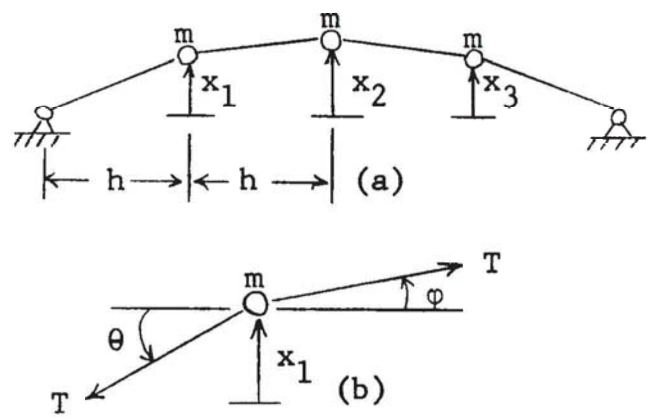


Figure 5.7

$$m \frac{d^2 x_1}{dt^2} = - T \sin \theta + T \sin \varphi.$$

For small displacements $\sin \theta \cong \tan \theta$, $\sin \varphi \cong \tan \varphi$ so that

$$m \frac{d^2 x_1}{dt^2} = - T \tan \theta + T \tan \varphi = T \left\{ - \frac{x_1}{h} + \frac{x_2 - x_1}{h} \right\} .$$

(a) In a similar way, derive the other two equations of motion and show that the characteristic equation leads to eigenvalues $-2\omega^2$, $(-2-\sqrt{2})\omega^2$, $(-2+\sqrt{2})\,\omega^2$, where $\omega^2 = T/(mh)$.

(b) Find, and draw diagrams of, the three normal mode configurations. Determine the natural frequencies.

5.6 Assume that in the drilling problem of Example 5.1 a calm sea prevails, so that the amplitude of the platform oscillation is zero, $U = 0$. However, the platform is now given a sudden constant downward displacement U_0 which results from an helicopter landing on the deck.

(a) Show that a particular solution due to this form of excitation is

$$x_p = \begin{pmatrix} k+K \\ K \end{pmatrix} \frac{U_0}{(k+2K)} ,$$

and construct the general solution.

(b) If the two lengths of pipe are at rest when the helicopter lands, determine the four equations for the four integration constants c_1, c_2, c_3 and c_4 of equation (5.28). Show that these equations are satisfied by the relations $c_1 = c_2 = - U_0/4$, $c_3 = c_4 = - U_0 k/4(k+2K)$, and thus determine the response of the system. Write the result in real form. Answer:

$$\begin{pmatrix} x \\ y \end{pmatrix} = -\begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{U_0}{2} \cos \omega_0 t - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{U_0 k}{2(k+2K)} \cos \omega_0 s t + \begin{pmatrix} k+K \\ k \end{pmatrix} \frac{U_0}{k+2K} .$$

$$\text{Answer: } x = \begin{pmatrix} k+K \\ K \end{pmatrix} \frac{U_0}{(k+2K)} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{U_0}{2} \cos \omega_0 t$$

$$+ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \frac{U_0 k}{2(k+2K)} \cos \omega_0 st.$$

5.7 Figure 5.8 is a top view of a two-dimensional version of the elastic structure of Problem 5.5. Particles of mass m at each of the four interior points, connected to the square rigid frame by the twelve equal strings of tension T, make small vertical oscillations about their equilibrium positions.

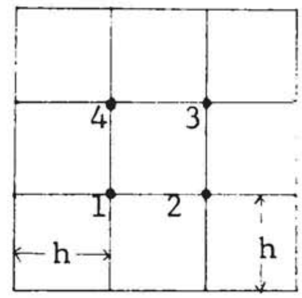


Figure 5.8

(a) Using the approximations of Problem 5.5, show that the equations of motion are $\frac{d^2x}{dt^2} = \omega^2 Ax$, where $\omega^2 = T/(mh)$ and

$$A = \begin{pmatrix} -4 & 1 & 0 & 1 \\ 1 & -4 & 1 & 0 \\ 0 & 1 & -4 & 1 \\ 1 & 0 & 1 & -4 \end{pmatrix}$$

(b) Show that A has the eigenvalues -2, -4, -4, -6, with an associated set of independent eigenvectors

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}.$$

(c) Write the general solution for the motion of this system.

(d) Describe the modes of vibration corresponding to each of the four partial solutions obtained from (b).

(e) Show that $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$ are also independent eigenvectors associated with the eigenvalue $\lambda = -4$, and describe the corresponding modes of vibration.

5.8 If v_1, v_2, v_3 are independent eigenvectors associated with the eigenvalues $\alpha^2 (\alpha \neq 0), 0, 0$, prove that the six solutions

$$v_1 e^{\alpha t} \ , \ v_2 \ , \ v_3 \ , \ v_1 e^{-\alpha t} \ , \ v_2 t \ , \ v_3 t \ ,$$

are independent.

5.9 Many problems involving small oscillations lead to a system of equations of the form

$$B \frac{d^2 x}{dt^2} = Ax \ ,$$

where A and B are symmetric matrices and B is positive definite. (See, for example, Karman and Biot, <u>Mathematical Methods in Engineering</u>, McGraw-Hill Book Co., New York, 1940, Chapter VI.) From the results of Chapter 5, Section 9, show that such a system can always be solved by the methods of this section as modified by Problem 1.8. Show also that the solutions are bounded oscillations if and only if A is negative definite.

6. A Single Linear Differential Equation with Constant Coefficients.

In the previous sections we considered systems of simultaneous first or second order equations; however, in many situations it is more natural to deal with a single differential equation of higher order. The point is well illustrated by the equations for the simple harmonic oscillator,

$$m \frac{dv}{dt} = - ku ,$$
$$\frac{du}{dt} = v. \tag{6.1}$$

If instead of treating equations (6.1) as a pair of simultaneous first order equations, we eliminated one of the variables, then a single equation in one variable would result. Inspection of (6.1) shows that v can be easily eliminated by substituting $\frac{du}{dt}$ for v in the first equation. We obtain

$$m \frac{d^2u}{dt^2} = - ku. \tag{6.2}$$

Of course (6.2) is just Newton's law in the form mass times acceleration equals force. Thus the two first order equations have been replaced by one second order differential equation.

Equation (6.2) is just the special case, for n = 1, of the system (5.4) of the previous section. However, since we wish to generalize (6.2) in a different way, we shall repeat some of the steps in the solution of this equation.

Once again we seek a solution of the form

$$u = Ue^{\lambda t} \ , \ U \neq 0 \ , \tag{6.3}$$

by substituting this expression for u in (6.2). This gives

$$m\lambda^2 Ue^{\lambda t} = - kUe^{\lambda t} \ , \tag{6.4}$$

which is satisfied (identically) if and only if

$$\lambda^2 + \frac{k}{m} = 0. \tag{6.5}$$

If (6.5) is compared with (1.15) it is seen that it is nothing other than the characteristic equation. The roots of (6.5) are

$$\lambda_1 = i\omega \ , \ \lambda_2 = -i\omega, \text{ where } \omega^2 = k/m. \tag{6.6}$$

Consequently there are two solutions of the form (6.3),

$$u_1 = U_1 e^{i\omega t} \ , \ u_2 = U_2 e^{-i\omega t} \ , \tag{6.7}$$

and the general solution is

$$u = U_1 e^{i\omega t} + U_2 e^{-i\omega t}. \tag{6.8}$$

The constants U_1 and U_2 are determined from initial conditions which, as in previous sections, will consist of specified initial values for u and v. Thus, the initial conditions are of the form

$$u(t_0) = U_0 \ , \ v(t_0) = \left(\frac{du}{dt}\right)_{t_0} = V_0. \tag{6.9}$$

For example, if (with $t_0 = 0$) the mass is given an initial displacement U_0 but zero velocity, then $u(0) = U_0$ and $\left(\frac{du}{dt}\right)_0 = 0$. with these initial conditions solution (6.8) gives the relations $U_1 = U_2 = U_0/2$, and substitution of these values into (6.8) yields

$$u = U_0 \frac{e^{i\omega t}+e^{-i\omega t}}{2} = U_0 \cos \omega t. \tag{6.10}$$

As expected, solution (6.10) is identical to that previously obtained by solving the system of two first order equations; see (1.25) or (1.26). This example illustrates the connection between the two methods of approach and outlines some essential features in the solution of a single higher order linear differential equation with constant coefficients.

Homogeneous Linear Second Order Equations with Constant Coefficients

Consider the second order homogeneous equation

$$\frac{d^2x}{dt^2} + a_1 \frac{dx}{dt} + a_0 x = 0 , \tag{6.11}$$

in which a_1 and a_0 are constants. This can be reduced to a system of first order equations by defining $\frac{dx}{dt}$ as a new variable y so that $\frac{d^2x}{dt^2} = \frac{dy}{dt}$ and we get

$$\begin{aligned} \frac{dy}{dt} &= - a_1 y - a_0 x , \\ \frac{dx}{dt} &= y. \end{aligned} \tag{6.12}$$

Equations (6.12) are a special case of (1.1); thus we expect the solution to be of the form $x = Xe^{\lambda t}$ and $y = Ye^{\lambda t}$. Substitution in (6.12) gives the characteristic equation

$$(6.13) \qquad \det(\lambda I-A) = \begin{vmatrix} \lambda+a_1 & a_0 \\ -1 & \lambda \end{vmatrix} = \lambda^2 + a_1\lambda + a_0 = 0 ,$$

which determines the eigenvalues

$$(6.14) \qquad \lambda_1 = \frac{1}{2}\left[-a_1 + \sqrt{a_1^{\,2}-4a_0}\right] , \quad \lambda_2 = \frac{1}{2}\left[-a_1 - \sqrt{a_1^{\,2}-4a_0}\right] .$$

The initial conditions which accompany (6.12) specify $x(t_0)$ and $y(t_0)$, so the appropriate initial conditions for (6.11) are specified values of $x(t_0)$ and $\left(\frac{dx}{dt}\right)_{t_0}$.

Rather than work with systems of equations, we may wish to solve equation (6.11) directly. Knowing that the solution is of exponential form we let

$$(6.15) \qquad x = Xe^{\lambda t} , \ X \neq 0.$$

Substitution into (6.11) gives

$$(\lambda^2 + a_1\lambda + a_0)Xe^{\lambda t} = 0.$$

This equation is satisfied if and only if

$$(6.16) \qquad \lambda^2 + a_1\lambda + a_0 = 0.$$

Equation (6.16) is identical with the characteristic equation (6.13). Assuming two distinct roots λ_1,λ_2 we have two solutions of the form (6.15) which when added give the general solution

$$x = X_1 e^{\lambda_1 t} + X_2 e^{\lambda_2 t}. \tag{6.17}$$

The constants X_1 and X_2 are determined by the initial values $x(t_0)$ and $\left(\frac{dx}{dt}\right)_{t_0}$.

Example 6.1. Determine the solution to the differential equation

$$\frac{d^2x}{dt^2} + 2\,\frac{dx}{dt} + 2x = 0. \tag{6.18}$$

A solution of the form $x = Xe^{\lambda t}$ gives the characteristic equation

$$\lambda^2 + 2\lambda + 2 = 0.$$

The roots of this equation are

$$\lambda = -1 \pm \sqrt{1-2} \quad \text{or} \quad \lambda_1 = -1+i \ , \ \lambda_2 = -1-i \ ,$$

so that the solutions

$$X_1 e^{(-1+i)t} \quad , \quad X_2 e^{-(1+i)t} \ ,$$

when combined, give the general solution of (6.18) as

$$x = X_1 e^{(-1+i)t} + X_2 e^{-(1+i)t}. \tag{6.19}$$

An alternative form of writing the solution is obtained by using the Euler relations

$$e^{\pm it} = \cos t \pm i \sin t.$$

The solution now takes the form

$$x = e^{-t}\{X_1(\cos t + i \sin t) + X_2(\cos t - i \sin t)\} ,$$

which becomes

$$x = e^{-t}\{X_1' \cos t + X_2' \sin t\} , \tag{6.20}$$

where the new constants X_1' and X_2' are defined in terms of X_1 and X_2 by the relations

$$X_1' = X_1+X_2 , \; X_2' = i(X_1-X_2). \tag{6.21}$$

If initial conditions are given, say $x(0) = 1$ and $\left(\frac{dx}{dt}\right)_0 = 0$, we can find X_1 and X_2 from (6.19), thus:

$$1 = X_1 + X_2 ,$$

$$0 = (-1+i)X_1 - (1+i)X_2 ,$$

from which

$$X_1 = \tfrac{1}{2}(1-i) , \; X_2 = \tfrac{1}{2}(1+i).$$

These can be substituted in (6.19) to give the solution in complex exponential form. However if we want the solution in real form (as we usually do) we can proceed directly to (6.20). Here the initial conditions give

$$1 = X_1' ,$$

$$0 = X_1' - X_2',$$

and the solution is therefore

$$x = e^{-t}(\cos t + \sin t).$$

Example 6.2. In Example 2.2, analysis of the resistor-capacitor network of Figure 2.1 led to the coupled voltage equations (2.34),

$$\text{(2.34)}\qquad \begin{aligned} \frac{dV_1}{dt} &= -\frac{2}{3RC}V_1 - \frac{1}{3RC}V_2 , \\ \frac{dV_2}{dt} &= -\frac{1}{3RC}V_1 - \frac{2}{3RC}V_2 . \end{aligned}$$

The voltage V_2 can be eliminated by solving for V_2 from the first equation of (2.34) and substituting this value in the second equation. This leads to the following second order equation

$$\text{(6.22)}\qquad \frac{d^2V_1}{dt^2} + \frac{4}{3RC}\frac{dV_1}{dt} + \frac{1}{3(RC)^2}V_1 = 0.$$

Initial conditions $V_1(0) = V_0, V_2(0) = 0$, when substituted in the first of equations (2.34), give an initial condition on $\frac{dV_1}{dt}$. This condition, together with the initial condition $V_1(0) = V_0$, are clearly equivalent to the original set of initial conditions, so we may state the initial conditions for (6.22) as

$$\text{(6.23)}\qquad V_1(0) = V_0 , \quad \left(\frac{dV_1}{dt}\right)_{t=0} = -\frac{2}{3RC}V_0 .$$

If we assume a solution of the form $V_1 = Ee^{\lambda t}$, substitution in (6.22) gives the characteristic equation

$$\lambda^2 + \frac{4}{3RC}\lambda + \frac{1}{3(RC)^2} = \left(\lambda + \frac{1}{3RC}\right)\left(\lambda + \frac{1}{RC}\right) = 0.$$

Summing the two solutions corresponding to the two values of λ, we obtain

(6.24) $$V_1 = E_1 e^{-\frac{1}{3RC}t} + E_2 e^{-\frac{1}{RC}t}.$$

Using the initial conditions (6.23) in (6.24) we get

$$V_0 = E_1 + E_2 ,$$

$$-\frac{2}{3RC} V_0 = -\frac{1}{3RC} E_1 - \frac{1}{RC} E_2 ,$$

which give the constants E_1 and E_2 as

$$E_1 = E_2 = V_0/2.$$

The solution (6.24) thus becomes

$$V_1 = \frac{V_0}{2} e^{-\frac{1}{3RC}t} + \frac{V_0}{2} e^{-\frac{1}{RC}t}.$$

This result is the same as that previously obtained; see equations (2.44).

We have assumed that λ_1 and λ_2 are distinct so that two independent solutions are obtained. If the discriminant in (6.15) vanishes, then $\lambda_1 = \lambda_2$, and the two solutions are not independent. Under this condition of a multiple root we have but one solution of the form $ce^{\lambda t}$ and another independent solution must be found. Thus, we are given the second order equation

(6.11) $$\frac{d^2x}{dt^2} + a_1 \frac{dx}{dt} + a_0 x = 0 ,$$

with $a_1{}^2 - 4a_0 = 0$, and one solution

(6.15) $$x = X_1 e^{\lambda_1 t} ,$$

where

(6.25) $$\lambda_1 = -a_1/2.$$

Take a second solution in the form of an unknown function v(t) multiplied by the known solution (6.17). We have

(6.26) $$x_1 = v(t)e^{\lambda_1 t},$$

so that

$$\frac{dx_1}{dt} = v'e^{\lambda_1 t} + v\lambda_1 e^{\lambda_1 t} ,$$

$$\frac{d^2x_1}{dt^2} = v''e^{\lambda_1 t} + 2\lambda_1 v'e^{\lambda_1 t} + v\lambda_1{}^2 e^{\lambda_1 t}.$$

If these values are substituted into (6.12) we obtain

(6.27) $$\{v'' + [2\lambda_1+a_1]v' + [\lambda_1{}^2+a_1\lambda_1+a_0]v\}e^{\lambda_1 t} = 0.$$

The coefficient of the v term is zero because it is just the left side of the characteristic equation with multiple root λ_1. In addition, the coefficient of the v' term is zero because of (6.25). Equation (6.27) thus reduces to

(6.28) $$v'' = 0.$$

Integrating twice with respect to t gives

(6.29) $$v = C_2 t + C_1 ,$$

where C_2 and C_1 are constants. Using this equation for v, the solution (6.26) becomes

$$x_1 = (C_2t+C_1)e^{\lambda_1 t} = C_2te^{\lambda_1 t} + C_1e^{\lambda_1 t}. \tag{6.30}$$

If this second solution (6.30) is compared with the first solution (6.15), we see that (6.30) contains the new part $C_2te^{\lambda_1 t}$ as well as the original solution $C_1e^{\lambda_1 t}$. Hence, we conclude that a second order differential equation having a characteristic equation with double roots λ_1 has one solution $X_1e^{\lambda_1 t}$ and a second solution $X_2te^{\lambda_1 t}$ so that the total solution is

$$x_1 = X_1e^{\lambda_1 t} + X_2te^{\lambda_1 t}. \tag{6.31}$$

Example 6.3. Find the solution of the differential equation

$$\frac{d^2x}{dt^2} - 4\frac{dx}{dt} + 4x = 0.$$

Assuming a solution $x = X_1e^{\lambda t}$ gives the characteristic equation

$$\lambda^2-4\lambda+4 = (\lambda-2)(\lambda-2) = 0 ,$$

which has the double root $\lambda_1 = \lambda_2 = 2$. One solution is X_1e^{2t} and a second solution is X_2te^{2t}. The total solution is

$$x = X_1e^{2t} + X_2te^{2t}.$$

Problems

6.1 (a) Solve the following set of simultaneous equations

$$\frac{dx_1}{dt} = x_1 - 2x_2 ,$$

$$\frac{dx_2}{dt} = x_1 - 2x_2 ,$$

with initial conditions $x_1(0) = 3$, $x_2(0) = 1$, by getting one equation in x_2 and solving this differential equation. Evaluate all constants and also determine x_1.

Answer: See Problem 1.4.

(b) Solve the system of equations

$$\frac{dx}{dt} = - 3x + 2y$$

$$\frac{dy}{dt} = - 2y$$

with initial conditions $x(0) = 1$, $y(0) = 1$, by solving the one differential equation in x, evaluating constants, and also finding y.

Answer: See Problem 1.6.

6.2 (a) Given the differential equation

$$\frac{d^2x}{dt^2} - 2x = 0$$

with initial conditions $x(0) = 2$, $\left(\frac{dx}{dt}\right)_0 = 0$, solve the equation. Also determine a system of first order equations and corresponding initial conditions which replace the single given equation.

Answer: $x = 2 \cosh \sqrt{2}t$.

(b) Solve the differential equation

$$\frac{d^2x}{dt^2} + 4 \frac{dx}{dt} + 8x = 0$$

with initial conditions $x(0) = 3$, $\left(\frac{dx}{dt}\right)_0 = 0$.

Also determine a system of first order equations and corresponding initial conditions which replace the single given equation.

Answer: $x = 3e^{-2t}\{\cos 2t + \sin 2t\}$.

6.3 Find the solution to the following differential equations:

(a) $\frac{d^2x}{dt^2} - x = 0$, with $x(0) = 1$, $\left(\frac{dx}{dt}\right)_0 = 0$.

Answer: $x = \cosh t$.

(b) $\frac{d^2y}{dt^2} - 3 \frac{dy}{dt} + 2y = 0$.

(c) $\frac{d^2x}{dt^2} + 9x = 0$, with $x(0) = 3$, $\left(\frac{dx}{dt}\right)_0 = 0$.

Answer: $x = 3 \cos 3t$.

(d) $\frac{d^2x}{dt^2} - 2 \frac{dx}{dt} + 2x = 0$.

(e) $\frac{d^2x}{dt^2} + 6 \frac{dx}{dt} + 13x = 0$.

6.4 Find the solution to the following differential equations:

(a) $\frac{d^2x}{dt^2} + 2\frac{dx}{dt} + x = 0$, with $x(0) = 0$, $\left(\frac{dx}{dt}\right)_0 = 1$.

Answer: $x = te^{-t}$.

(b) $\frac{d^2y}{dt^2} + 6\frac{dy}{dt} + 9y = 0$.

(c) $4\frac{d^2x}{dt^2} - 4\frac{dx}{dt} + x = 0$, with $x(0) = 1$, $\left(\frac{dx}{dt}\right)_0 = 0$.

Answer: $x = (1 - \frac{t}{2})e^{t/2}$.

(d) $\frac{d^2x}{dt^2} = 0$.

(e) $\frac{d^2x}{dt^2} + 4\frac{dx}{dt} + 4x = 0$

6.5 (a) Find the solution of the differential equation

$$\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 5y = 0$$

which satisfies the initial conditions

$y = 0$, $\frac{dy}{dx} = 2$, when $x = 0$.

Express your answer in a form involving no complex numbers.

Answer: $y = 2e^{-2x}\sin x$.

(b) Find the solution of the equation

$$\frac{d^2y}{dx^2} + 4\,\frac{dy}{dx} + 4y = 0$$

which satisfies the initial conditions

$$y(0) = 1, \quad \left(\frac{dy}{dx}\right)_0 = 1.$$

Briefly discuss the behavior of the solution for large x and then for small x.

6.6 Figure 3.8 shows a mass m attached to a spring of stiffness k and to a viscous dashpot having coefficient c; omit the force F. If m = 1 and k = 16, it is desired to find how the response varies with different values of c. Assume an initial deflection x(0) = 15. Neglect gravity. Find the response and sketch the x-t curve for each of the cases: (a) c = 10. (b) c = 8. (c) c = 0.8.

6.7 Tanks A and B are connected as shown in Figure 6.1. When t = 0 tank A contains 100 gal. of brine having 50 lb. of salt in solution, while tank B contains 50 gal. of water. Water is run into tank A at a rate of $\frac{4}{3}$ gal./min. and brine from tank B is discharged at the rate of $\frac{4}{3}$ gal./min.

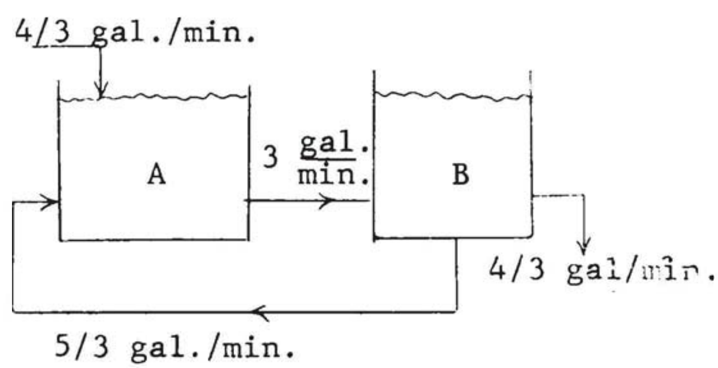


Figure 6.1

Mixture from B is returned to A at the rate of $\frac{5}{3}$ gal./min.

Find the system of differential equations for the amounts of salt in tanks A and B at any time. From these equations find a single differential equation and corresponding initial conditions for the amount of salt in tank A at any time.

6.8 A compound has two isomers A and B. A certain organism can digest the isomer B thus creating compound C, but it cannot digest isomer A. The reaction is symbolized by

$$A \underset{2}{\overset{1}{\rightleftharpoons}} B \xrightarrow{1} C$$

where the numbers are the rate constants for the transformations. Initially there is no B present and 3 moles of isomer A when a catalyst is added at time $t = 0$ to start the reaction. If $N_A(t)$ and $N_B(t)$ denote the number of moles of A and B present at time t, show that for $t \geqq 0$ the reaction obeys the equations

$$\frac{dN_A}{dt} = - N_A + 2N_B ,$$

$$\frac{dN_B}{dt} = N_A - 3N_B .$$

Determine the single differential equation and corresponding initial conditions for N_B. Solve this equation and hence find $N_B(t)$. Sketch the curve for N_B as a function of t.

Answer:

$$\frac{d^2N_B}{dt^2} + 4\,\frac{dN_B}{dt} + N_B = 0.$$

$$N_B = \sqrt{3}\; e^{-2t}\,\sinh\sqrt{3}\; t.$$

Homogeneous n^{th} Order Linear Equation with Constants Coefficients

Consider a differential equation of the form

$$(6.32)\qquad \frac{d^nx_1}{dt^n} + a_{n-1}\,\frac{d^{n-1}x_1}{dt^{n-1}} + \ldots + a_0x_1 = 0\ ,$$

in which the coefficients $a_{n-1},\ldots,a_0$ are constants. Equation (6.32) is said to be an n^{th} order equation because the highest derivative which occurs is $\frac{d^nx_1}{dt^n}$. We assume that in addition to equation (6.32) the following n initial values are given

$$(6.33)\qquad x_1(t_0),\ \left(\frac{dx_1}{dt}\right)_{t_0},\ldots,\ \left(\frac{d^{n-1}x_1}{dt^{n-1}}\right)_{t_0}.$$

As with a single second order equation, it is possible to express (6.32) in terms of a system of first order equations. Make the following definitions of new variables $x_2,x_3,\ldots,x_n$,

(6.34)
$$\frac{dx_1}{dt} = x_2 \ ,$$
$$\frac{dx_2}{dt} = x_3 \ , \text{ so that } \frac{d^2x_1}{dt^2} = x_3 \ ,$$
$$\frac{dx_3}{dt} = x_4 \ , \text{ so that } \frac{d^3x_1}{dt^3} = x_4 \ ,$$
$$\cdots\cdots\cdots\cdots$$
$$\frac{dx_{n-1}}{dt} = x_n, \text{ so that } \frac{d^{n-1}x_1}{dt^{n-1}} = x_n .$$

Definitions (6.34) together with equation (6.32) gives the following system of first order equations

(6.35)
$$\frac{dx_1}{dt} = x_2 \ ,$$
$$\frac{dx_2}{dt} = x_3 \ ,$$
$$\frac{dx_3}{dt} = x_4 \ ,$$
$$\cdots\cdots\cdots\cdots$$
$$\frac{dx_{n-1}}{dt} = x_n \ ,$$
$$\frac{dx_n}{dt} = -a_0x_1-a_1x_2-a_2x_3 \ \cdots \ -a_{n-1}x_n .$$

We know that there is a solution to the system (6.35) of the form

(6.36) $$x_1 = X_1e^{\lambda t} \ , \ x_2 = X_2e^{\lambda t} \ ,\ldots, \ x_n = X_ne^{\lambda t} \ ,$$

if and only if $\det(\lambda I-A) = 0$. In this case, $\det(\lambda I-A) = 0$ gives

$$(6.37) \qquad \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0$$

as the characteristic equation. The roots of (6.37) are the eigenvalues. If they are all distinct, they give n independent solutions of (6.35) of the form (6.36).

Rather than express (6.32) as the system (6.35), we adopt the viewpoint that we wish to solve (6.32) directly. Thus, we wish to solve an equation of the form

$$(6.38) \qquad \frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0 x = 0 ,$$

with given initial values

$$(6.39) \qquad x(t_0) , \left(\frac{dx}{dt}\right)_{t_0} , \ldots , \left(\frac{d^{n-1}x}{dt^{n-1}}\right)_{t_0} .$$

From equations (6.35) and their solution (6.36) we assume a solution of exponential form,

$$(6.40) \qquad x = Xe^{\lambda t} , \quad X \neq 0.$$

Substitution into (6.38) gives

$$(6.41) \qquad (\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0)Xe^{\lambda t} = 0,$$

which is satisfied if and only if

$$(6.42) \qquad \varphi(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0.$$

Equation (6.42) is the characteristic equation (6.37), and if it has n distinct roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ then n independent solutions of the form (6.40) exist. When added together these solutions give the general solution

$$x = X_1 e^{\lambda_1 t} + X_2 e^{\lambda_2 t} + \ldots + X_n e^{\lambda_n t}. \tag{6.43}$$

Solution (6.43) satisfies the n^{th} order differential equation 6.38), and the constants $X_1, X_2, \ldots, X_n$ are determined by the initial values (6.39).

<u>Example 6.4</u>. Determine the general solution of the differential equation

$$\frac{d^4 x}{dt^4} - \alpha^4 x = 0.$$

Letting

$$x = Xe^{\lambda t},$$

the characteristic equation is

$$\lambda^4 - \alpha^4 = (\lambda^2+\alpha^2)(\lambda^2-\alpha^2) = (\lambda+i\alpha)(\lambda-i\alpha)(\lambda+\alpha)(\lambda-\alpha) = 0.$$

The four roots are $\lambda_1 = -i\alpha$, $\lambda_2 = i\alpha$, $\lambda_3 = -\alpha$, $\lambda_4 = \alpha$ and the general solution is

$$x = X_1 e^{-i\alpha t} + X_2 e^{i\alpha t} + X_3 e^{-\alpha t} + X_4 e^{\alpha t},$$

or if we wish to avoid complex numbers

$$x = X_1' \sin\alpha t + X_2' \cos\alpha t + X_3 e^{-\alpha t} + X_4 e^{\alpha t}.$$

Example 6.5. Find the general solution of the differential equation

$$\frac{d^3y}{dx^3} = y.$$

Assuming a solution of the form $y = Ye^{\lambda x}$ we get

$$\lambda^3 = 1$$

as the characteristic equation. The values of λ are then the three cube roots of 1, namely

$$\lambda_1 = 1 ,$$

$$\lambda_2 = e^{2\pi i/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2} i ,$$

$$\lambda_3 = e^{4\pi i/3} = -\frac{1}{2} - \frac{\sqrt{3}}{2} i.$$

Consequently the general solution can be written as

$$y = Y_1 e^{-x} + Y_2 e^{(-\frac{1}{2} + \frac{\sqrt{3}}{2} i)x} + Y_3 e^{(-\frac{1}{2} - \frac{\sqrt{3}}{2} i)x}$$

or

$$y = Y_1 e^{-x} + e^{-x/2}(Y_2' \cos \frac{\sqrt{3}x}{2} + Y_3' \sin \frac{\sqrt{3}x}{2}).$$

If equation (6.42) does not have n distinct roots then (6.38) has fewer than n solutions of the form $Xe^{\lambda t}$, and to get the general solution we must find additional solutions. We shall show that if $\lambda = c$ is an r-fold root of (6.42) then (6.38) has the r solutions

$$(6.43) \qquad e^{ct},\ te^{ct},\ t^2e^{ct}, \ldots,\ t^{r-1}e^{ct}.$$

Since the sum of the multiplicities of the distinct roots of (6.42) is n, this will give the requisite n independent solutions of the differential equation.

We first observe that if the polynomial equation $\varphi(\lambda) = 0$ has c as an r-fold root then

$$(6.44) \qquad \varphi(c) = 0,\ \varphi'(c) = 0,\ \varphi''(c) = 0, \ldots, \varphi^{(r-1)}(c) = 0.$$

For in this case $\varphi(\lambda)$ has $(\lambda-c)^r$ as a factor, that is,

$$\varphi(\lambda) = (\lambda-c)^r\psi(\lambda).$$

Therefore

$$\varphi'(\lambda) = (\lambda-c)^r\psi'(\lambda) + r(\lambda-c)^{r-1}\psi(\lambda)$$

$$= (\lambda-c)^{r-1}[(\lambda-c)\psi'(\lambda) + r\psi(\lambda)]$$

has $(\lambda-c)^{r-1}$ as a factor. Using the same argument on $\varphi'(\lambda)$ we see that $\varphi''(\lambda)$ has $(\lambda-c)^{r-2}$ as a factor. Continuing in this way we eventually find that $\varphi^{(r-1)}(\lambda)$ has $\lambda-c$ as a factor. From this (6.44) follows, for any polynomial having a power of $\lambda-c$ as a factor must vanish when we set $\lambda = c$.

We have already seen that for any value of λ,

$$(6.45) \qquad \frac{d^n}{dt^n} e^{\lambda t} + a_{n-1} \frac{d^{n-1}}{dt^{n-1}} e^{\lambda t} + \ldots + a_1 \frac{d}{dt} e^{\lambda t} + a_0 e^{\lambda t}$$

$$= \varphi(\lambda)e^{\lambda t}.$$

This is an identity in λ and t; hence we can differentiate both sides with respect to λ. To see the effect of this on the left-hand side of (6.45) consider just one of the terms, for example $a_2 \frac{d^2}{dt^2} e^{\lambda t}$. Since we are considering both λ and t as variables it is more appropriate to write this as $a_2 \frac{\partial^2}{\partial t^2} e^{\lambda t}$. Then

$$\frac{\partial}{\partial\lambda} (a_2 \frac{\partial^2}{\partial t^2}) e^{\lambda t} = a_2 \frac{\partial}{\partial\lambda} (\frac{\partial^2}{\partial t^2} e^{\lambda t})$$

$$= a_2 \frac{\partial^2}{\partial t^2} (\frac{\partial}{\partial\lambda} e^{\lambda t}) = a_2 \frac{\partial^2}{\partial t^2} (te^{\lambda t}) ,$$

since we may interchange the order of differentiation with respect to t and λ. A similar argument will work with any of the terms, so we can say that the effect of differentiating the left-hand side of (6.45) with respect to λ is just to replace $e^{\lambda t}$ by $te^{\lambda t}$. Differentiating a second time will replace this in turn by $t^2 e^{\lambda t}$, and so on.

Turning to the right-hand side of (6.45), a differentiation with respect to λ gives

(6.46) $$[\varphi'(\lambda) + t\varphi(\lambda)]\, e^{\lambda t} ,$$

a second differentiation gives

(6.47) $$[\varphi''(\lambda) + 2t\varphi'(\lambda) + t^2\varphi(\lambda)]\, e^{\lambda t} ,$$

and in general a k-th differentiation gives an expression of the form

(6.48) $[\varphi^{(k)}\lambda + c_{k-1}t\varphi^{(k-1)}(\lambda) + \ldots + c_1 t^{k-1}\varphi'(\lambda) + t^k\varphi(\lambda)]e^{\lambda t}$.

Now let us put these pieces together. We assume that c is an r-fold root of $\varphi(\lambda) = 0$, so that the equations (6.44) hold. It follows that the right-hand side of (6.45) is zero when we set $\lambda = c$ and so e^{ct} is a solution of (6.38). (We already know this). Differentiating (6.45) with respect to λ gives, in light of the above discussion,

$$\frac{d^n}{dt^n}(te^{\lambda t}) + a_{n-1}\frac{d^{n-1}}{dt^{n-1}}(te^{\lambda t}) + \ldots + a_1\frac{d}{dt}(te^{-\lambda t}) + a_0(te^{-\lambda t})$$

$$= [\varphi'(\lambda) + t\varphi(\lambda)]e^{\lambda t}.$$

On setting $\lambda = c$ the right-hand side of this equation becomes zero by virtue of the first two of equations (6.44), and so te^{ct} is a solution of (6.38). If $r \geqq 3$, a second differentiation with respect to λ, plus the fact that by virtue of the first three of equations (6.44), the expression (6.47) is zero for $\lambda = c$, gives us t^2e^{ct} as a solution. Comparing (6.44) and (6.48) we see that the process can continue for r-1 differentiations, thus establishing the fact that

$$e^{ct},\ te^{ct},\ t^2e^{ct},\ldots,t^{r-1}e^{ct}$$

are solutions of the differential equation.

A linear combination of these r solutions can be written in the convenient form

$$(c_1 + c_2 t + \ldots + c_r t^{r-1})e^{ct} , \tag{6.49}$$

or, if we wish, in the form $P(t)e^{ct}$, where $P(t)$ is a polynomial of degree $r-1$. The solution of the most general equation of the form (6.38) can then be expressed as follows:

Let the equation

$$\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0$$

have the distinct roots $\lambda_1, \ldots, \lambda_k$ with respective multiplicities $r_1, \ldots, r_k$, where each each $r_i \geqq 1$.
Let $P_1(t), \ldots, P_k(t)$ be any polynomials of degrees $r_1-1, \ldots, r_k-1$ respectively. Then

$$x(t) = P_1(t)e^{\lambda_1 t} + \ldots + P_k(t)e^{\lambda_k t} \tag{6.50}$$

is a solution of the equation

$$\frac{d^x n}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0 x = 0. \tag{6.51}$$

To show that (6.50) is the general solution of (6.51) it is necessary to prove that the set of n solutions of the form $t^j e^{\lambda_i t}$ is linearly independent. For such a proof we refer the reader to W. S. Kaplan, <u>Ordinary Differential Equations</u>, Addison-Wesley Publishing Co., Reading, Mass. 1958, Chapter 4, Section 8; or to E. A. Coddington, <u>Ordinary Differential Equations</u>, Prentice-Hall, Englewood Cliffs, New Jersey, 1961, Chapter 2.

Example 6.6. Determine the general solution to the equation

$$\frac{d^4x}{dt^4} + 2\,\frac{d^3x}{dt^3} + \frac{d^2x}{dt^2} = 0.$$

If $x = Xe^{\lambda t}$ then the characteristic equation is

$$\lambda^4 + 2\lambda^3 + \lambda^2 = \lambda^2(\lambda^2 + 2\lambda + 1) = \lambda^2(\lambda+1)^2 = 0.$$

There is a double root $\lambda_1 = \lambda_2 = 0$ and another double root $\lambda_3 = \lambda_4 = -1$. From (6.49), the solution corresponding to roots $\lambda_1 = \lambda_2 = 0$ is

$$x = (c_1+c_2t)e^{0t} = c_1+c_2t\ ,$$

and the solution for roots $\lambda_3 = \lambda_4 = -1$ is

$$x = (d_1+d_2t)e^{-t}\ ,$$

so that the general solution is

$$x = c_1+c_2t+(d_1+d_2t)e^{-t}.$$

Problems

6.9 (a) Solve the differential equation

$$\frac{d^3y}{dt^3} - 3\,\frac{dy}{dt} + 2y = 0$$

subject to the initial conditions $y(0) = 0$, $(\frac{dy}{dt})_0 = 1$, $(\frac{d^2y}{dt^2})_0 = 2$. Answer: $y = te^t$.

(b) Replace the given differential equation by an equivalent system of three first order equations and corresponding initial conditions. Show that the characteristic equation is the same as in part (a).

6.10 For the system of three equations of Problem 2.2 with initial conditions $x_1(0) = 1$, $x_2(0) = 1$, $x_3(0) = 2$;

(a) Show by elimination that the single differential equation in x_1 is

$$\frac{d^3x_1}{dt^3} - 3\frac{d^2x_1}{dt^2} - 6\frac{dx_1}{dt} + 8x_1 = 0.$$

(Hint: use x_3 from the first equation in each of the second and third equations to get two equations involving x_1 and x_2. From these two equations get one equation in x_1.)

(b) Solve this single equation in x_1 and determine the integration constants from suitable initial conditions.

6.11 Find the solutions to the following differential equations:

(a) $\frac{d^3x}{dt^3} + 3\frac{d^2x}{dt^2} - \frac{dx}{dt} - 3x = 0$ with $x(0) = 1$, $(\frac{dx}{dt})_0 = 0$, $(\frac{d^2x}{dt^2})_0 = 5$. Answer: $x = e^t + \frac{1}{2}e^{-3t} - \frac{1}{2}e^{-t}$.

(b) $\frac{d^3x}{dt^3} + 3\frac{d^2x}{dt^2} - 6\frac{dx}{dt} - 8x = 0.$

(c) $\frac{d^3y}{dt^3} + y = 0.$

Answer: $y = Y_1 e^{-t} + e^{\frac{t}{2}}(Y_2 \cos \frac{\sqrt{3}t}{2} + Y_3 \sin \frac{\sqrt{3}t}{2}).$

(d) $\frac{d^3y}{dt^3} - y = 0$, with $y(0) = 1$, $\left(\frac{dy}{dt}\right)_0 = \left(\frac{d^2y}{dt^2}\right)_0 = 0.$

6.12 Find the solutions to the following differential equations:

(a) $\frac{d^4y}{dt^4} - 2\frac{d^2y}{dt^2} + y = 0.$

Answer: $y = (c_1+c_2t)e^t + (d_1+d_2t)e^{-t}.$

(b) $\frac{d^4y}{dt^4} + 2\frac{d^2y}{dt^2} + y = 0$ with $y(0) = 0$, $(\frac{dy}{dt})_0 = 1,$

$(\frac{d^2y}{dt^2})_0 = 2$, $(\frac{d^3y}{dt^3})_0 = 1.$

Answer: $y = 2 \sin t - t \cos t + t \sin t.$

(c) $\frac{d^3x}{dt^3} - 3\frac{d^2x}{dt^2} + 3\frac{dx}{dt} - x = 0.$

6.13 In Example 5.2, the Mohole problem, the system was idealized to that shown in Figure 5.2 for the free oscillations. Using Newton's law in the form F = ma show that the governing differential equations are

$$m\frac{d^2x_1}{dt^2} = -k(x_1-u) - K(x_1-x_2),$$

$$m \frac{d^2x_2}{dt^2} = - kx_2 - K(x_2 - x_1) ,$$

which on eliminating x_2 give the single equation

$$\frac{d^4x_1}{dt^4} + \frac{2(k+K)}{m} \frac{d^2x_1}{dt^2} + \frac{k(k+2K)}{m^2} x_1 = 0.$$

Show that the solution of this fourth order equation leads to the characteristic equation (5.17). Hence, determine the solution $x_1(t)$. Describe how you would find $x_2(t)$ but do not actually solve for it.

6.14 For the network of Figure 6.2 write the differential equations for the currents in each loop and combine them to obtain a single equation for the current I_2 through the resistor R.

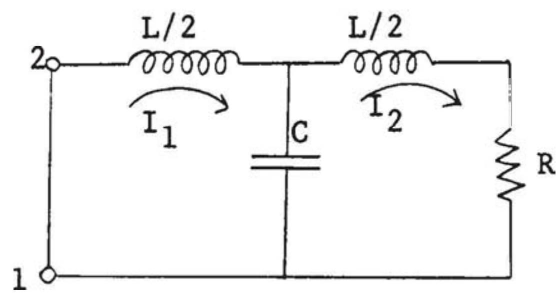


Figure 6.2

6.15 For the coupled two loop L-C circuit of Problem 5.2 (Figure 5.5) determine the fourth-order equation in either Q_1 or Q_2. Find the natural frequencies of oscillation and the corresponding normal mode configurations.

6.16 In the three tank Problem 2.8 (Figure 2.8) determine a single differential equation for the concentration, C_2, in Tank 2 at any time. Determine concentration C_2 when the concentration in Tank 1 is $C_0/10$.

Nonhomogeneous n^{th} Order Equation

A differential equation of the form (6.38), but with a function g(t) on the right-hand side is called an n^{th} order linear nonhomogeneous equation with constant coefficients.

$$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0 x = g(t). \tag{6.52}$$

Together with the n initial values

$$x(t_0)\ ,\ \left(\frac{dx}{dt}\right)_{t_0},\ldots,\ \left(\frac{d^{n-1}x}{dt^{n-1}}\right)_{t_0} \tag{6.53}$$

the determination of a solution is called the initial value problem. From our previous experience with a system of non-homogeneous equations we are led to consider a solution in two parts. First determine the general solution to the homogeneous equation

$$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0 x = 0.$$

This solution, x_c, is called the complementary solution of (6.52). Then find a particular solution, x_p, of (6.52). The general solution of (6.52) is then

$$x = x_c + x_p\ , \tag{6.54}$$

and the given initial values (6.53) can now be used to determine the n constants that appear in x_c.

If $g(t)$ is expressible as a sum of functions,

$$g(t) = g_1(t) + \ldots + g_m(t)$$

the superposition principle can be used to determine x_p as a sum

$$x_p = x_{p1} + \ldots + x_{pm} ,$$

where x_{pi}, for $i = 1,\ldots,m$, is a solution of (6.52) with $g(t)$ replaced by $g_i(t)$.

The method of finding x_c is known from previous work. Again we are faced with having to determine a particular integral x_p; again we consider a restricted class of functions $g(t)$ and use the method of undetermined coefficients in order to construct a solution.

We know from previous work that sometimes it is convenient to replace $g(t)$ by a new function which is complex-valued. For such cases, the proof of Theorem 4.4 can be readily modified to show that if

$$x_p = x_1(t) + ix_2(t)$$

is a solution of (6.52) with $g(t)$ replaced by $g_1(t)+ig_2(t)$ then $x_1(t)$ and $x_2(t)$ are solutions with $g(t)$ replaced by $g_1(t)$ and $g_2(t)$ respectively, assuming, of course, that g_1, g_2, x_1, and x_2 are real functions of t.

We now look for a particular solution when $g(t)$ is constant, exponential, or sinusoidal.

<u>When $g(t)$ is constant</u> equation (6.52) becomes

(6.55) $$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0 x = G.$$

Inspection shows that if we take x_p constant, $x_p = c_1$, then all the derivatives vanish leaving

$$a_0 c_1 = G \quad \text{or} \quad c_1 = G/a_0 ,$$

provided $a_0 \neq 0$. The case $a_0 = 0$ will be considered later.

<u>When g(t) is exponential</u> equation (6.52) is

(6.56) $$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0 x = Ge^{pt}.$$

We have seen that a function of the form

(6.57) $$x_p = ce^{pt}$$

when substituted in (6.56) gives

(6.58) $$\varphi(p)ce^{pt} = Ge^{pt} ,$$

where, following the notation of (6.42) we have put

$$\varphi(p) = p^n + a_{n-1}p^{n-1} + \ldots + a_1 p + a_0.$$

Equation (6.57) is satisfied if

(6.59) $$c = \frac{G}{\varphi(p)} ,$$

provided $\varphi(p) \neq 0$. Since the roots of $\varphi(\lambda) = 0$ are just the λ_i, we see that a particular solution of the form (6.57) can be found provided $p \neq \lambda_i$. If $p = \lambda_i$ then $\lambda(p) = 0$ and equation (6.56) cannot be satisfied by a function of the form ce^{pt}. We will investigate later how to find a particular solution when $p = \lambda_i$.

<u>When g(t) is sinusoidal</u> equation (6.52) has the form

$$\text{(6.60)} \qquad \frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0 x = G \sin pt .$$

If we replace G sin pt by Ge^{ipt} then we can assume a solution in the form

$$\text{(6.61)} \qquad x_p = ce^{ipt} ,$$

so that we get from the previous case

$$\text{(6.62)} \qquad c = G/\varphi(ip) ,$$

provided $ip \neq \lambda_i$, one of the roots of the characteristic equation. The constant c can be written in polar form

$$c = re^{i\theta}.$$

A particular solution with $g(t) = Ge^{ipt}$ is given by

$$x_p = re^{i(pt+\theta)} ,$$

and the imaginary part of this solution gives

$$x_p = r \sin(pt+\theta)$$

as a particular solution of (6.60).

We must now investigate how to determine a particular solution when the form of g(t) is such that the apparent form of the particular solution, x_p, coincides with one of the complementary solutions. The difficulty arises if we are led to assume a particular solution which is of the type

$$x_p = ce^{\lambda_i t}.$$

When this is substituted into the differential equation in order to determine c, the result is zero and not g(t), for x_p is a solution of the homogeneous equation. The case when $g(t) = G \neq 0$ and $a_0 = 0$ is included in this statement of the problem.

The difficulty can be resolved by the procedure given below. Given the differential equation (6.52)

$$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0 x = g(t);$$

we recall that if $\lambda = p$ is a root of the characteristic equation (6.42)

$$\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0$$

of multiplicity $r(\geq 1)$ then the complementary solution contains the terms

$$e^{pt}\ ,\ te^{pt}\ ,\ t^2e^{pt}\ ,\ldots,\ t^{r-1}e^{pt}.$$

It then turns out that if

$$g(t) = P_h(t)e^{pt}\ , \tag{6.63}$$

where $P_h(t)$ is a polynomial of degree $h(\geq 0)$, then a corresponding particular solution is of the form

$$x_p = Q_h(t)t^r e^{pt}\ , \tag{6.64}$$

where $Q_h(t)$ is also a polynomial of degree h. That is

(6.65) $$Q_h = c_0+c_1 t +\dots+ c_h t^h.$$

The constants $c_0,\dots,c_h$ can be found by the method of undetermined coefficients.

This procedure works equally well if $r = 0$, that is, if p is not a root of the characteristic equation. In this case it generalizes the three special cases considered above, since they fall under (6.64) with $r = 0$, $h = 0$ and p either zero, real, or complex.

For a different way of looking at this procedure and a proof of its validity the reader is referred to Coddington, loc. cit.

Example 6.7. Solve the equation

$$\frac{d^2x}{dt^2} - 2\,\frac{dx}{dt} + x = 4+e^t\ .$$

The characteristic equation is

$$\lambda^2 - 2\lambda + 1 = (\lambda-1)^2 = 0.$$

Thus, $\lambda_1 = \lambda_2 = 1$ and the complementary solution is

$$x_c = X_1 e^t+X_2 te^t.$$

With $g_1(t) = 4$ we can see by inspection that $x_{p1} = 4$ is a particular solution. With $g_2(t) = e^t$ we observe that X_1e^t already appears in the complementary solution. Using the procedure previously given we have $p = 1$, $h = 0$ in (6.63). Further, $\lambda = 1$

is a double root of the characteristic equation so that r = 2. Hence, from (6.64) we take

$$x_{p2} = At^2e^t$$

and substitute in the differential equation

$$\frac{d^2x}{dt^2} - 2\,\frac{dx}{dt} + x = e^t$$

to find the constant A; we get

$$\left[2Ae^t+4Ate^t+At^2e^t\right] -2\left[2Ate^t+At^2e^t\right] +At^2e^t = e^t.$$

All terms on the left cancel except $2Ae^t$ so that

$$2Ae^t = e^t.$$

Hence A = 1/2. Using the superposition principle the general solution is

$$x = X_1e^t+X_2te^t+4 + \frac{1}{2}\,t^2e^t.$$

<u>Example 6.8</u>. In Example 3.2 where the simple harmonic oscillator was excited by the force $F_0\sin pt$ we found that if p is close to ω then the amplitude of the particular solution becomes very large and we have the condition of resonance. However, if $p = \omega$ then a particular integral of the form $c \sin \omega t$ cannot be found. We are given the differential equation

$$\frac{d^2x}{dt^2} + \omega^2x = \frac{F_0}{m}\sin \omega t. \tag{6.66}$$

Replace $\frac{F_0}{m}$ sin ωt by $\frac{F_0}{m}\, e^{i\omega t}$ and the differential equation becomes

(6.67) $$\frac{d^2x}{dt^2} + \omega^2 x = \frac{F_0}{m}\, e^{i\omega t}.$$

From (6.9) the complementary solution is

(6.68) $$x_c = X_1 e^{i\omega t} + X_2 e^{-i\omega t}.$$

In equation (6.67), $g(t) = \frac{F_0}{m} e^{i\omega t}$ and from past work we would initially seek a particular solution of the form $x_p = ce^{i\omega t}$. However, $ce^{i\omega t}$ is contained in complementary solution (6.68), so we use the procedure previously described. Comparing $g(t) = \frac{F_0}{m}\, e^{i\omega t}$ with (6.63) we see that $p = i\omega$ and $h = 0$. Further, the root $\lambda = i\omega$ is of multiplicity 1, so that $r = 1$. Hence we adopt a particular solution of the form

(6.69) $$x_p = Ate^{i\omega t}.$$

The undetermined coefficient A is found by substituting (6.69) into (6.67); we get

$$\left[2i\omega Ae^{i\omega t} - \omega^2 Ate^{i\omega t}\right] + \omega^2 \left[Ate^{i\omega t}\right] = \frac{F_0}{m}\, e^{i\omega t} ,$$

or

$$2i\omega Ae^{i\omega t} = \frac{F_0}{m}\, e^{i\omega t}.$$

Equating coefficients we obtain

$$A = \frac{-F_0 i}{2m\omega} ,$$

and the particular solution (6.69) becomes

(6.70) $$x_p = \frac{-F_0 i}{2m\omega} t e^{i\omega t}.$$

In order to determine a particular solution of (6.66) we take the imaginary part of (6.70) and get

(6.71) $$x_p = \frac{-F_0}{2m\omega} t \cos \omega t.$$

We conclude that when the excitation frequency is equal to the natural frequency, $p = \omega$, then the amplitude of the particular solution is proportional to t. Mathematically, the amplitude increases indefinitely; physically, the system would eventually fail or cease to be linear.

Problems

6.17 Determine the general solution to the following equations,

(a) $\frac{d^4x}{dt^4} - 2\frac{d^2x}{dt^2} + x = 3 \sin t + 4.$

Answer: $x = X_1e^t + X_2e^{-t} + X_3te^t + X_4te^{-t} + 4 + \frac{3}{4} \sin t.$

(b) $\frac{d^3x}{dt^3} - 4\frac{dx}{dt} = 3e^t + \cos 2t.$

(c) $\frac{d^3x}{dt^3} - \frac{d^2x}{dt^2} + \frac{dx}{dt} - x = t + 5e^{2t}.$

Answer: $x = X_1e^t + X_2e^{it} + X_2e^{-it} + e^{2t} - t - 1.$

6.18 Solve the equations:

(a) $\frac{d^3x}{dt^3} - 2\,\frac{d^2x}{dt^2} + \frac{dx}{dt} = 4 + 2e^{2t}.$

Answer: $x = X_1+X_2e^t+X_3te^t+4t+e^{2t}.$

(b) $\frac{d^2y}{dt^2} - y = t^2 + 3.$

Answer: $y = Y_1e^t+Y_2e^{-t}-t^2-5.$

(c) $\frac{d^4x}{dt^4} + 3\,\frac{d^2x}{dt^2} - 4x = 40e^t+18\cos t.$

Answer: $x = X_1e^t+X_2e^{-t}+X_3e^{2it}+X_4e^{-2it}+4te^t-3\cos t.$

6.19 Find the general solutions to the following equations.

(a) $\frac{d^2y}{dt^2} + y = \sin 3t.$

(b) $\frac{d^2y}{dt^2} + y = \sin t.$

(c) $\frac{d^3y}{dt^3} - \frac{dy}{dt} = 4e^{-t}+3e^{2t}$ with $y(0) = 0$, $\left(\frac{dy}{dt}\right)_0 = -1$,

$\left(\frac{d^2y}{dt^2}\right)_0 = 2.$

6.20 For each of the following equations find the complementary solution and the form of the particular solution but do not actually evaluate the undetermined coefficients.

(a) $\frac{d^2y}{dt^2} + 2\frac{dy}{dt} + 5y = 8e^{-t} + e^{-t}\sin 2t.$

(b) $\frac{d^2x}{dt^2} + 4\frac{dx}{dt} + 4x = (t^2 + 1)e^{-2t}.$

(c) $\frac{d^4x}{dt^4} + 2\frac{d^3x}{dt^3} + \frac{d^2x}{dt^2} = 2 + 3t + t^2e^{-t}.$

6.21 Consider the space station problem, 5.4, shown in Figure 5.6.

(a) From the equations of Problem 5.4(a) find the one differential equation governing the motion of the mass M. (Hint: let $x_2+x_3 = u$ and add the second two equations.)

(b) If the mass M has an oscillating force $F_0\sin pt$ acting on it, determine the governing equation of motion of the mass M. Find a particular solution and determine if resonance is possible.

6.22 The circuit of Figure 6.2 is called a low-pass filter since it transmits low frequencies well, but attenuates high frequencies. Find the steady state voltage, I_2R, across the resistor R when a sinusoidal input voltage of unit amplitude and frequency (a) 200 cps and (b) 2000 cps is applied across terminals 1-2, if the circuit constants are L = 0.6 henries, c = 0.16 μf, and R = 2000 ohms.

6.23 The circuit of Figure 6.3 is called a high-pass filter since it transmits high frequencies well, but attenuates low frequencies. Find the steady-state voltage, I_2R across the resistor R when a sinusoidal input voltage of unit amplitude and frequency (a) 300 cps and (b) 3000 cps is applied across the terminals, if the circuit constants are L = 0.15 henries, C = 0.04 μf, R = 2000 ohms.

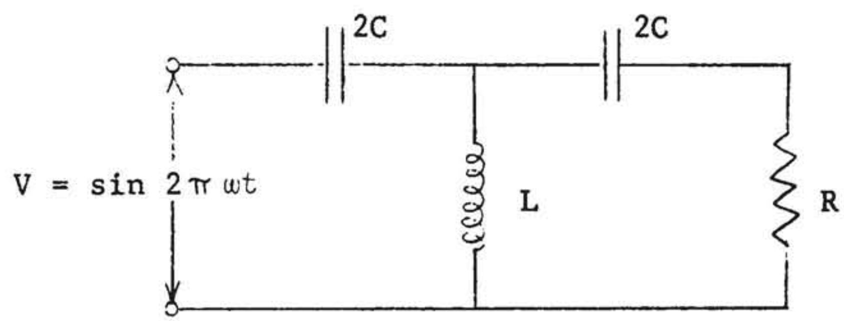


Figure 6.3

Answer: $$RI_2 = \left[\left(1 - \frac{1}{2CLp^2}\right) - \frac{1}{CRp}\left(1 - \frac{1}{4CLp^2}\right) i\right]^{-1} e^{ipt},$$

where $p = 2\pi w$.

6.24 A mass is suspended from a hook by a rubber band whose unstretched length is one foot. When allowed to hang at rest the mass is two feet below the hook. If the mass is held at the hook and dropped how low does it go? Assuming no friction and neglecting the mass of the rubber band describe the subsequent motion of the mass and graph its position as a function of time.

Partial Answer: The motion is periodic with period $(\pi + 2 \arctan (1/\sqrt{2}) + 2)/\sqrt{g} = 1.13$ sec.

6.25 Equal masses are suspended by equal springs and a cord over a pulley of negligible mass. There is friction at the pulley which opposes an angular velocity θ' of the pulley by a torque $h\theta'$.

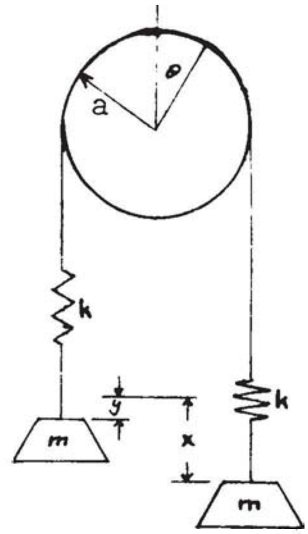


Figure 6.4

Set up equations of motion, find the eigenvalues, and discuss the motion.

Comments and partial answers:

(1) If x, y and θ are displacements from equilibrium as indicated in Figure 6.4, and $u = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix}$, the equations of motion can be written in the form

$$\begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & 0 \end{pmatrix} \frac{d^2u}{dt^2} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & h \end{pmatrix} \frac{du}{dt} + k \begin{pmatrix} 1 & 0 & -a \\ 0 & 1 & a \\ -a & a & 2a^2 \end{pmatrix} u = 0.$$

(2) If $u = Ue^{\lambda t}$, where U is a constant vector, then λ satisfies

$$\begin{vmatrix} m\lambda^2+k & 0 & -ak \\ 0 & m\lambda^2+k & ak \\ -ak & ak & h\lambda+2a^2k \end{vmatrix} = 0.$$

(3) The eigenvalues are $0, \pm i\omega, -c \pm \sqrt{c^2-\omega^2}$, where $\omega^2 = k/m$, $c = a^2k/h$.

CHAPTER 7

Vector Analysis in 3-Space

1. Vector Algebra in 3-Space

We recall from Chapters 2 and 5 the basic notion of a vector space with an inner product over the real numbers. In this chapter we will be concerned mainly with the 3-dimensional space V_3 with its usual inner product, defined by

$$(u_1,u_2,u_3)\cdot(v_1,v_2,v_3) = u_1v_1 + u_2v_2 + u_3v_3 . \tag{1.1}$$

By setting up in 3-dimensional Euclidean space E_3 a fixed Cartesian coordinate system we associate triples of real numbers with points of E_3 and thus give a geometric interpretation to the algebraic operations in V_3. We will not always distinguish between a point of E_3 and the vector in V_3 given by its coordinates. We will also sometimes think of the vector (u_1,u_2,u_3) as represented by the directed line segment from the origin 0 to the point A with coordinates (u_1,u_2,u_3).

The inner product (1.1) has the following geometrical interpretation. If $\vec{u} = (u_1,u_2,u_3)$ is represented by 0A and $\vec{v} = (v_1,v_2,v_3)$ is represented by 0B, as in Figure 1.1, then

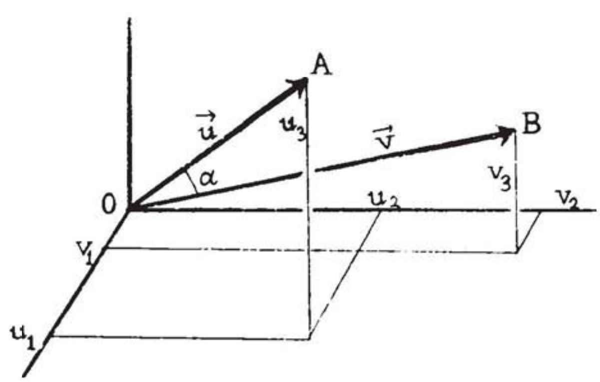


Figure 1.1

$$\vec{u}\cdot\vec{v} = \|\vec{u}\|\ \|\vec{v}\| \cos\alpha \ , \tag{1.2}$$

where α is the angle AOB. Of course $\|\vec{u}\|$, $\|\vec{v}\|$ are just the lengths OA, OB; and we could have taken (1.2) as our definition of the inner product and then deduced formula (1.1) from geometrical considerations. Formula (1.1) is the more useful in calculations and proofs, while (1.2) is the more useful for intuitive understanding and geometrical visualization. Thus the crucial relation

$$\vec{u}\cdot(a\vec{v} + b\vec{w}) = a(\vec{u}\cdot\vec{v}) + b(\vec{u}\cdot\vec{w}), \qquad a,b \text{ scalars}, \tag{1.3}$$

is easy to prove using (1.1); while the fact that the equality $\vec{u}\cdot\vec{v} = 0$ expresses the perpendicularity of $\vec{u}$ and $\vec{v}$ is immediately clear from (1.2).

We now introduce a new product in V_3, namely the <u>cross product</u> or <u>vector product</u>. This is a rule assigning to a pair of vectors $\vec{u},\vec{v}$ in V_3 a vector $\vec{u}x\vec{v}$ in V_3 by the formula

$$(u_1,u_2,u_3)x(v_1,v_2,v_3) = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1). \tag{1.4}$$

<u>Note</u>. The concept of cyclic permutation is useful in remembering formulas relating to cross products. If the digits 1,2,3 are equally spaced around a circle (Figure 1.2) a rotation through 120° carries 1 into 2, 2 into 3, and 3 into 1. This change we call a <u>cyclic permutation</u> C of the digits. There

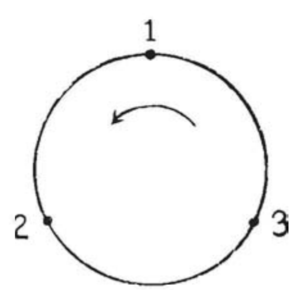


Figure 1.2

is one other cyclic permutation, carrying 1 into 3, 3 into 2, and 2 into 1, which may be considered to be either the inverse of C or the result of applying C twice.

By applying this concept to equation (1.4) we need merely remember that the 1-st component of the cross product is $u_{\underline{2}}v_{\underline{3}} - u_{\underline{3}}v_{\underline{2}}$. Cyclic permutations of the five underlined digits then give the 2-nd component to be $u_{\underline{3}}v_{\underline{1}} - u_{\underline{1}}v_{\underline{3}}$, and the 3-rd component to be $u_{\underline{1}}v_{\underline{2}} - u_{\underline{2}}v_{\underline{1}}$.

We now state some properties of the cross product which indicate its geometrical significance.

Theorem 1.1. The vector $\vec{u}x\vec{v}$ is orthogonal to $\vec{u}$ and to $\vec{v}$, and is zero if $\vec{u},\vec{v}$ are dependent.

We leave the proof as an easy exercise for the reader.

Theorem 1.2. $\|\vec{u}x\vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \sin\alpha$, where α is the angle between $\vec{u}$ and $\vec{v}$, $0 \leqq \alpha \leqq \pi$.

Proof. Schwarz's Inequality (Theorem 2.3 of Chapter 5) tells us that $\|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u}\cdot\vec{v})^2 \geqq 0$. For vectors in V_3 we can prove the more precise statement

$$\|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u}\cdot\vec{v})^2 = \|\vec{u}x\vec{v}\|^2 \tag{1.5}$$

by multiplying out and verifying the identity

$$(u_1{}^2 + u_2{}^2 + u_3{}^2)(v_1{}^2 + v_2{}^2 + v_3{}^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2$$

$$= (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2 .$$

But $(\vec{u}\cdot\vec{v})^2 = \|\vec{u}\|^2\|\vec{v}\|^2 \cos^2\alpha$ and $1 - \cos^2\alpha = \sin^2\alpha$. Thus

$$|\vec{u}\times\vec{v}|^2 = |\vec{u}|^2|\vec{v}|^2 \sin^2\alpha ,$$

and the theorem is proved, noting that $\sin\alpha \geqq 0$ if $0 \leqq \alpha \leqq \pi$.

Theorems 1.1 and 1.2 give an almost complete geometric interpretation to the cross product. If $\vec{u},\vec{v}$ are dependent (that is, collinear) then $\vec{u}\times\vec{v} = 0$. Otherwise, let π be the plane determined by $\vec{u},\vec{v}$. Then Theorem 1.1 says that $\vec{u}\times\vec{v}$ is in the direction normal to π and Theorem 1.2 tells us the magnitude of $\vec{u}\times\vec{v}$. However this leaves an ambiguity, as there are two (equal and opposite) vectors normal to π and with a given magnitude. To resolve the ambiguity we restrict the chosen coordinate system by insisting that it be <u>right-handed</u>; that is, the axes OX, OY, OZ are to be such that turning from OX to OY would drive a right-handed screw up OZ (see Figure 1.3). Then the cross product $\vec{u}\times\vec{v}$ may be characterized as follows.

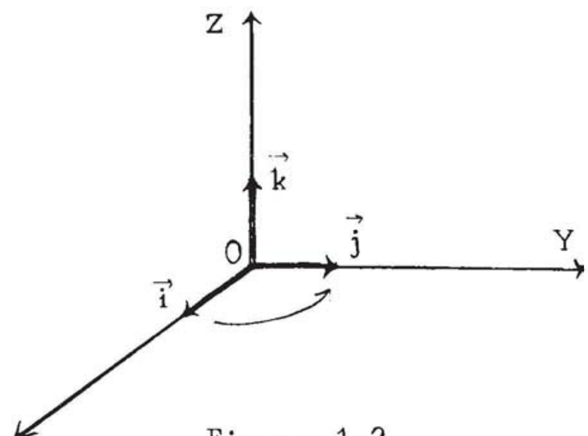


Figure 1.3

Choose a rotation in π bringing the line of action of $\vec{u}$ into coincidence with that of $\vec{v}$ and let α be the angle of the rotation. Let $\vec{n}$ be the unit vector along the normal representing the movement of a <u>right-handed</u> screw turned in π in the direction of the rotation. Then

$$\vec{u}\times\vec{v} = \|\vec{u}\|\ \|\vec{v}\|\ (\sin\alpha)\ \vec{n}\ . \tag{1.6}$$

It is perfectly possible to develop the theory of cross-products taking (1.6) as <u>definition</u>. If we then introduced a coordinate system into E_3 we would find the cross product given in component form by (1.4), provided the coordinate system had been chosen to be right-handed. Note that, by contrast with the cross product, the geometric interpretation of the <u>inner</u> product did not depend on the right-handedness of the coordinate system.

We now develop the algebraic properties of the cross product. As usual, the symbols $\vec{i},\vec{j},\vec{k}$ stand for the unit vectors along the 3 coordinate axes; that is, for the natural basis in V_3.

<u>Theorem 1.4</u>.

$$\begin{aligned} \vec{i}\times\vec{j} &= \vec{k}, & \vec{j}\times\vec{i} &= -\vec{k},\\ \vec{j}\times\vec{k} &= \vec{i}, & \vec{k}\times\vec{j} &= -\vec{i},\\ \vec{k}\times\vec{i} &= \vec{j}, & \vec{i}\times\vec{k} &= -\vec{j}\ . \end{aligned}$$

<u>Note</u>. By spacing $\vec{i},\vec{j},\vec{k}$ around our circle in Figure 1.2 we can apply cyclic permutation to get the second and third rows of Theorem 1.4 from the first row.

The reader should notice that the <u>order</u> in which two vectors are multiplied is important, $\vec{i}\times\vec{j} \neq \vec{j}\times\vec{i}$, etc. In fact, reference to (1.4) immediately shows that the situation indicated in Theorem 1.4 is quite general, namely

Theorem 1.5. The cross product is anti-commutative: $\vec{u}x\vec{v} = -\vec{v}x\vec{u}$.

The cross product behaves very well with respect to the vector space structure in V_3. Thus we have

Theorem 1.6.

$$\vec{u}x(\vec{v}+\vec{w}) = \vec{u}x\vec{v}+\vec{u}x\vec{w}, \quad (\vec{u}+\vec{v})x\vec{w} = \vec{u}x\vec{w}+\vec{v}x\vec{w} ;$$

$$a(\vec{u}x\vec{v}) = (a\vec{u})x\vec{v} = \vec{u}x(a\vec{v}) \quad \text{where a is a scalar .}$$

The proof is an easy exercise based on (1.4); once again we see that to prove a property of the cross product it is better to use the algebraic rather than the geometric description (1.6).

We may express Theorem 1.6 by saying that premultiplication by $\vec{u}$ and postmultiplication by $\vec{u}$ are both linear maps of V_3 into itself.

The algebraic discussion of the cross product will be complete when we have studied the relation between the inner product and the cross product, and when we have investigated the validity of the associative law for cross products. (In fact, it turns out that $\vec{u}x(\vec{v}x\vec{w})$ and $(\vec{u}x\vec{v})x\vec{w}$ are not equal!)

It is important to keep in mind that the inner product of two vectors is a scalar, while the cross product is a vector. Thus the expression $(\vec{u}\cdot\vec{v})x\vec{w}$ has no meaning, while the expression $\vec{u}\cdot(\vec{v}x\vec{w})$ does have a meaning and represents a scalar. This implies that the symbol u·vxw is, in fact, unambiguous; nevertheless we will insert the parentheses in the interest of clarity. Now let $\vec{u} = (u_1,u_2,u_3)$, $\vec{v} = (v_1,v_2,v_3)$, $\vec{w} = (w_1,w_2,w_3)$.

<u>Theorem 1.7</u>.

$$\vec{u}\cdot(\vec{v}\times\vec{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = (\vec{u}\times\vec{v})\cdot\vec{w} .$$

The proof is a simple computation, using, of course, (1.4). We then find that $\vec{u}\cdot(\vec{v}\times\vec{w})$ is evaluated as the expansion of the determinant by its first row, and $(\vec{u}\times\vec{v})\cdot\vec{w}$ as the expansion by the last row. The interchangeability of the dot and the cross in this triple product tempts us to use a more symmetrical notation. One in common use is $[\vec{u},\vec{v},\vec{w}] = (\vec{u}\times\vec{v})\cdot\vec{w} = \vec{u}\cdot(\vec{v}\times\vec{w})$. Since $[\vec{u},\vec{v},\vec{w}]$ is a scalar it is called the <u>triple scalar product</u>.

From the properties of determinants we immediately infer

<u>Theorem 1.8</u>.

$$[\vec{u},\vec{v},\vec{w}] = [\vec{v},\vec{w},\vec{u}] = [\vec{w},\vec{u},\vec{v}]$$
$$= -[\vec{u},\vec{w},\vec{v}] = -[\vec{v},\vec{u},\vec{w}] = -[\vec{w},\vec{v},\vec{u}] .$$

<u>Note</u>. Here we have a somewhat different application of cyclic permutation. The first line of Theorem 1.8 can be stated as follows: The value of a triple scalar product is unchanged by cyclic permutation of the factors. The relation $[\vec{u},\vec{v},\vec{w}] = -[\vec{u},\vec{w},\vec{v}]$ states that the triple scalar product is changed in sign if we permute two of its factors.

We may assign an important geometric significance to the absolute value of the scalar $\vec{u}\cdot(\vec{v}\times\vec{w})$. Let $\vec{u},\vec{v},\vec{w}$

be represented by the lines OA, OB, OC as in Figure 1.4.

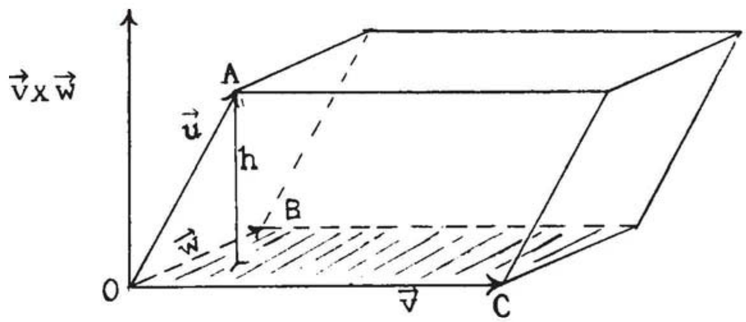


Figure 1.4

Suppose O, B, C not collinear (since, if they were, we would have $\vec{v}\times\vec{w} = \vec{0}$, $\vec{u}\cdot(\vec{v}\times\vec{w}) = 0$). Then $\vec{v}\times\vec{w}$ is represented by a line through 0 normal to the plane OBC, whose length is equal to the area of the parallelogram having OB, OC as sides. Thus the absolute value of $\vec{u}\cdot(\vec{v}\times\vec{w})$ is the product of the area of the parallelogram by the length of the projection of AO onto the normal to the plane of the parallelogram. But this length is just the <u>height</u> of the parallelepiped generated by the lines OA, OB, OC, when it is regarded as standing on the plane OBC. Observing finally that the product of the area of the base of a parallelepiped by its height is equal to its volume, we infer that <u>the absolute value of $\vec{u}\cdot(\vec{v}\times\vec{w})$ is the volume of the parallelepiped generated by OA, OB, OC</u>.

The question of when $\vec{u}\cdot(\vec{v}\times\vec{w}) = 0$ may now be studied either geometrically or algebraically. Approaching it from our geometrical interpretation of $\vec{u}\cdot(\vec{v}\times\vec{w})$, we see that $\vec{u}\cdot(\vec{v}\times\vec{w}) = 0$ if and only if the vectors $\vec{u},\vec{v},\vec{w}$ are coplanar. From our algebraic theory of matrices and determinants we know that the determinant of the matrix

$$\begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}$$

is zero if and only if the rows of the matrix are linearly dependent. Thus the geometric and algebraic stories are perfectly consistent.

<u>Example 1.1</u>. Find the volume of the tetrahedron whose vertices are at the points

$$(1,2,1),\ (2,-3,2),\ (1,1,4),\ (-1,2,2)\ .$$

First we may shift the origin to $(1,2,1)$ obtaining the points O,A,B,C with coordinates

$$(0,0,0),\ (1,-5,1),\ (0,-1,3),\ (-2,0,1)\ .$$

Next we observe that the volume of the tetrahedron is just 1/6 of the volume of the parallelpiped generated by OA, OB, OC. Thus the required volume is the absolute value of

$$1/6 \begin{vmatrix} 1 & -5 & 1 \\ 0 & -1 & 3 \\ -2 & 0 & 1 \end{vmatrix} = 9/2\ .$$

We turn now to the triple cross product, $\vec{u}\mathrm{x}(\vec{v}\mathrm{x}\vec{w})$. Since $\vec{v}\mathrm{x}\vec{w}$ is perpendicular to $\vec{v}$ and to $\vec{w}$, and since $\vec{u}\mathrm{x}(\vec{v}\mathrm{x}\vec{w})$ is perpendicular to $\vec{u}$ and to $\vec{v}\mathrm{x}\vec{w}$, it follows that $\vec{u}\mathrm{x}(\vec{v}\mathrm{x}\vec{w})$ <u>is a vector</u>

<u>in the plane of $\vec{v}$ and $\vec{w}$, perpendicular to $\vec{u}$</u>. This immediately shows that we cannot expect that $\vec{u}x(\vec{v}x\vec{w})$ and $(\vec{u}x\vec{v})x\vec{w}$ be equal, since the former lies in the plane of $\vec{v}$ and $\vec{w}$, while the latter lies in the plane of $\vec{u}$ and $\vec{v}$. In fact, we have already said enough to show that

$$\vec{u}x(\vec{v}x\vec{w}) = a\vec{v} + b\vec{w} ,$$

where a and b are scalars such that

$$a(\vec{u}\cdot\vec{v}) + b(\vec{u}\cdot\vec{w}) = 0.$$

In other terms, $\vec{u}x(\vec{v}x\vec{w}) = \lambda[(\vec{u}\cdot\vec{w})\vec{v}-(\vec{u}\cdot\vec{v})\vec{w}]$, for some scalar λ. This simple geometric line of reasoning does not enable us to determine λ easily. However, it does allow us to determine λ by looking at just one component on each side. Thus the first component of $\vec{u}x(\vec{v}x\vec{w})$ is evidently $u_2(v_1w_2-v_2w_1)-u_3(v_3w_1-v_1w_3)$.

The first component of $(\vec{u}\cdot\vec{w})\vec{v} - (\vec{u}\cdot\vec{v})\vec{w}$ is

$$(u_1w_1+u_2w_2+u_3w_3)v_1 - (u_1v_1+u_2v_2 + u_3v_3)\, w_1,$$

which, on cancelling the $u_1v_1w_1$ terms and regrouping, reduces to the same form. Hence $\lambda = 1$.

<u>Theorem 1.9</u>.

$$\vec{u}x(\vec{v}x\vec{w}) = (\vec{u}\cdot\vec{w})\vec{v} - (\vec{u}\cdot\vec{v})\vec{w};$$
$$(\vec{u}x\vec{v})x\vec{w} = (\vec{u}\cdot\vec{w})\vec{v} - (\vec{v}\cdot\vec{w})\vec{u}.$$

<u>Proof</u>. The first line we have proved above. The second follows from the first and the anti-commutativity of the cross product, thus:

$$(\vec{u}\times\vec{v})\times\vec{w} = -\vec{w}\times(\vec{u}\times\vec{v}) = \vec{w}\times(\vec{v}\times\vec{u})$$
$$= (\vec{w}\cdot\vec{u})\vec{v} - (\vec{w}\cdot\vec{v})\vec{u}$$
$$= (\vec{u}\cdot\vec{w})\vec{v} - (\vec{v}\cdot\vec{w})\vec{u}.$$

This is called the "Rule of the Middle Factor," since in either form the "middle factor" $\vec{v}$, of the left hand side appears first (with a scalar multiple) on the right hand side.

The ideas of coordinate geometry of 3 dimensions are often conveniently expressed in vector notation. Consider for example the straight line LL' in the direction of the vector $\vec{u}$ and passing through the point A corresponding to the vector $\vec{a}$. If P is any point on the line then (see Figure 1.5) $\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{OQ}$ and $\overrightarrow{OQ} = h\vec{u}$ for some scalar h. Thus if

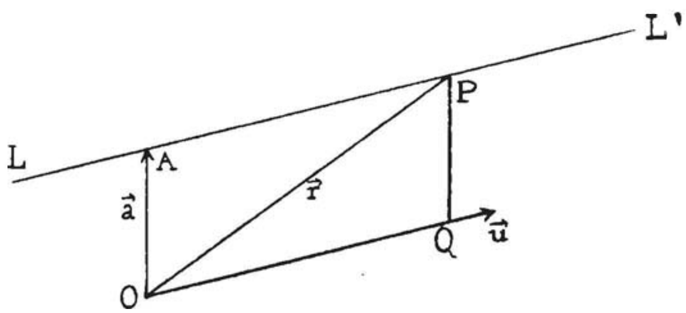


Figure 1.5

$\vec{r}$ is the position vector of P,

(1.7) $$\vec{r} = \vec{a} + h\vec{u} .$$

Conversely if $\vec{r}$ is given by (1.7), for some h, and if $\vec{r}$ represents $\overrightarrow{OP}$ then P lies on the given line. Thus (1.7) is the vector equation of the line in terms of a parameter h. If

$\vec{u}$ is a unit vector, then $|h|$ is the length of AP and h is positive if $\vec{u}$, with its origin translated to A, is directed towards P. If (1.7) is broken up into components with

$$\vec{r} = (x,y,z)$$
$$\vec{a} = (a_1,a_2,a_3),$$
$$\vec{u} = (u_1,u_2,u_3) ,$$

we obtain non-parametric equations of the line,

$$\frac{x-a_1}{u_1} = \frac{y-a_2}{u_2} = \frac{z-a_3}{u_3} . \tag{1.8}$$

provided u_1, u_2 and u_3 are all different from zero. The quantities (u_1,u_2,u_3) are called <u>direction</u> <u>ratios</u> of the line; if $\|\vec{u}\| = 1$ they are <u>direction</u> <u>cosines</u>.

Now consider a plane W containing the point A and the two distinct directions $\vec{u}$ and $\vec{v}$ (Figure 1.6). The point P, with

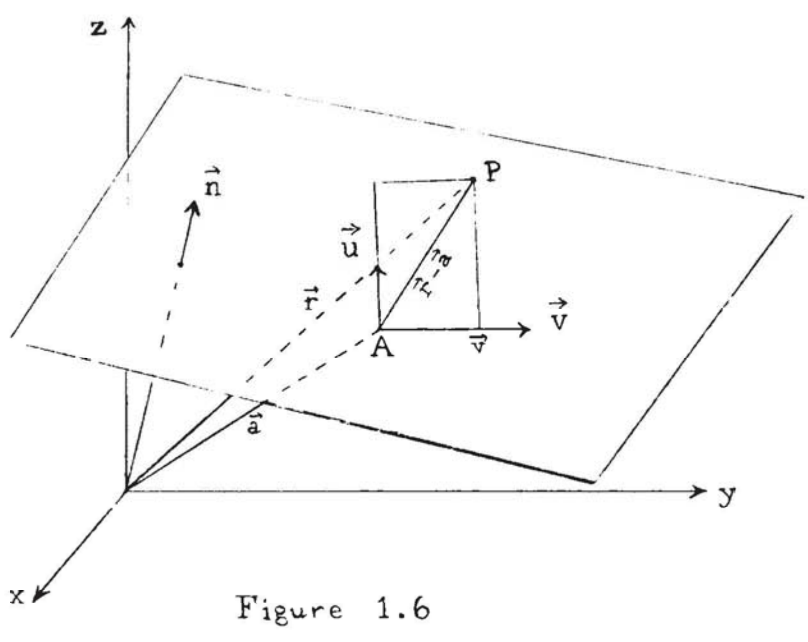


Figure 1.6

position vector $\vec{r}$, lies in W if and only if the three vectors $\vec{r}-\vec{a}$, $\vec{u}$, $\vec{v}$ are linearly dependent. Since $\vec{u},\vec{v}$ are not linearly dependent, this is equivalent to asking for scalars h,k such that

(1.9) $$\vec{r}-\vec{a} = h\vec{u} + k\vec{v} .$$

Thus (1.9) is the parametric vector equation of the plane, h and k being parameters.

A non-parametric vector equation of the plane W can be obtained in terms of a vector $\vec{n}$ perpendicular, or <u>normal</u>, to W. Since $\vec{n}$ is orthogonal to every vector lying in W we have $\vec{n}\cdot(\vec{r}-\vec{a}) = 0$. If we let C be the constant $\vec{n}\cdot\vec{a}$ this equation can be written as

(1.10) $$\vec{n}\cdot\vec{r} = C$$

Conversely, given any equation (1.10) with $\vec{n}$ a non-zero vector, if we define $\vec{a} = (C/\|\vec{n}\|^2)\vec{n}$ the equation can be written as $\vec{n}\cdot(\vec{r}-\vec{a}) = 0$, and hence is the equation of a plane.

Another field of application of vector concepts is mechanics, particularly the theory of motion about a fixed point 0, which we naturally take to be the origin. If a point P is rotating with angular velocity ω about an axis OA, (Figure 1.7) its velocity vector $\vec{v}$ satisfies

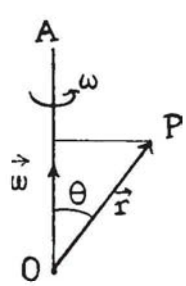


Figure 1.7

$$\|\vec{v}\| = \omega\|\vec{r}\|\sin\theta \ .$$

Taking into account the direction of $\vec{v}$ we see that if we introduce the vector angular velocity $\vec{\omega}$ acting along the axis of rotation we have

(1.11) $$\vec{v} = \vec{\omega}x\vec{r}.$$

Let a force $\vec{F}$ act at a point P, where $\overrightarrow{OP} = \vec{r}$. $\vec{F}$ tends to produce rotation about an axis through 0 perpendicular to the plane of $\vec{r}$ and $\vec{F}$. The moment of $\vec{F}$ about this axis is $L = \|\vec{F}\| \ \|\vec{r}\|\sin\theta$, and it tends to turn $\vec{r}$ in the direction towards $\vec{F}$. Hence we can introduce the vector moment of $\vec{F}$ about the point 0 by

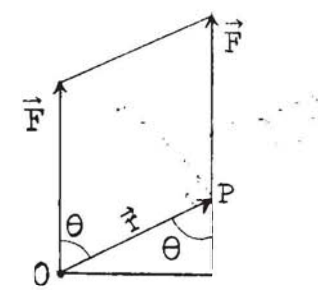


Figure 1.8

(1.12) $$\vec{L} = \vec{r}x\vec{F} \ .$$

In a similar fashion we define the moment of momentum, or angular momentum, of a particle of mass m and velocity $\vec{v}$, about the point 0, by

(1.13) $\vec{H} = \vec{r}x(m\vec{v}) = m\vec{r}x\vec{v}$.

By considering a body as the limit of a sum of particles of mass $\rho\Delta V$, where ρ is density and ΔV an element of volume, we can

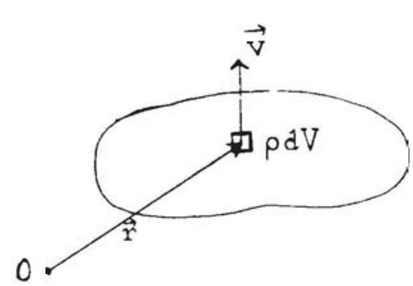


Figure 1.9

express the moment of momentum of the body, about 0, as

$$\vec{H} = \int_V \vec{r} x \vec{v} \rho \, dV.$$

Example 1.2. Angular Momentum of Conical Pendulum.

A conical pendulum consists of a slender rod whose end is fixed at 0 in such a way that it can be rotated with constant angular velocity ω about the vertical while at a fixed angle α from the vertical. It is desired to compute the moment of momentum of the rod with respect to point 0.

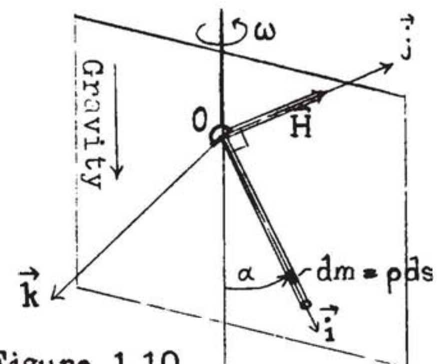


Figure 1.10

Introduce orthonormal vectors $\vec{i}, \vec{j}, \vec{k}$ with $\vec{i}$ along the rod and $\vec{j}$ in the plane of the rod and the vertical as shown in Figure 1.10. Thus, the basis vectors rotate with the rod with angular velocity ω. The rod has length L and mass M so that its mass per unit length is $\rho = M/L$. An element of mass $dm = \rho ds$ at position $\vec{r} = s\vec{i}$ moves on a circle of radius $s \sin\alpha$ with velocity $\omega s \sin\alpha(-\vec{k})$. The moment of momentum of the element about point 0 is

$$(s\vec{i})x(\rho ds \omega s \sin\alpha(-\vec{k})) = \rho\omega \sin\alpha \; s^2 (ds)\vec{j}.$$

The total moment of momentum of the rod is

$$\vec{H} = \int_0^L \rho\omega \sin\alpha \; s^2 ds \vec{j} = \vec{j}\rho\omega \sin\alpha \int_0^L s^2 ds$$

or

$$\vec{H} = \frac{(\rho L)L^2\omega \sin\alpha}{3}\,\vec{j} = \frac{ML^2\omega \sin\alpha}{3}\,\vec{j}.$$

As the rod rotates moving on a conical surface, the $\vec{H}$ vector also sweeps out a conical surface.

Problems

1.1 Prove Theorem 1.4 directly.

1.2 Prove Theorem 1.5.

1.3 Prove Theorem 1.6. What is the rank of the linear map $\vec{v} \to \vec{u}x\vec{v}$?

1.4 Deduce (1.4) from Theorems 1.4 and 1.6.

1.5 Calculate the volume of the tetrahedron with vertices at

(i) (0,0,0), (3,1,3), (4,1,4), (-1,2,1),

(ii) (2,1,4), (1,2,1), (0,0,0), (-3,3,-3),

(iii) (1,2,1), (-3,2,5), (3,2,5), (1,1,3) .

1.6 Prove that $\vec{u}x(\vec{v}x\vec{w}) + \vec{v}x(\vec{w}x\vec{u}) + \vec{w}x(\vec{u}x\vec{v}) = \vec{0}$.

[This is known as the <u>Jacobi identity</u> for vector products.]

1.7 Prove that if $\vec{u}$ is a unit vector then $\vec{u}x(\vec{v}x\vec{u})$ is the component of $\vec{v}$ orthogonal to $\vec{u}$.

1.8 Prove that

$$(\vec{a}x\vec{b})x(\vec{c}x\vec{d}) = [\vec{a},\vec{b},\vec{d}]\vec{c} - [\vec{a},\vec{b},\vec{c}]\vec{d}.$$

Prove also that $(\vec{a}x\vec{b})x(\vec{c}x\vec{d}) = (\vec{c}x\vec{d})x(\vec{b}x\vec{a})$, and hence obtain a linear relation between the vectors $\vec{a},\vec{b},\vec{c},\vec{d}$.

1.9 Show that the cross product (1.4) can be written as the determinant

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} .$$

1.10 Show that if (1.9) and (1.10) represent the same plane W then $\vec{u}$ and $\vec{v}$ are each orthogonal to $\vec{n}$. Hence show that

(a) If the equation of a plane is given in the form (1.9), equation in the form (1.10) can be obtained by taking for $\vec{n}$ the vector $\vec{u}\times\vec{v}$; and, in fact, the equation of the plane is $[\vec{r}-\vec{a},\vec{u},\vec{v}] = 0$.

(b) If the equation of a plane is given in the form (1.10) its equation in the form (1.9) can be obtained by taking for $\vec{u}$ and $\vec{v}$ any two independent solutions $\vec{w}_1$ and $\vec{w}_2$ of $\vec{n}\cdot\vec{w} = 0$.

1.11 Use the results of Problem 1.10 to find

(a) The equation of the plane through (1,-1,0) parallel to both the lines

$\frac{x-1}{2} = \frac{y+2}{3} = \frac{z-1}{-1}$, $\frac{x}{1} = \frac{y-1}{-1} = \frac{z+1}{2}$. Ans. $x-y-z = 2$.

(b) The parametric equation of the plane $x+2y-z = 4$.

1.12 Show that equation (1.7) is equivalent to $\vec{u}\times(\vec{r}-\vec{a}) = 0$.

1.13 Find the line of intersection of the two planes $x-y+z = 3$, $2x+y-2z = 5$. Ans. $\vec{r} = (\frac{8}{3}, -\frac{1}{3}, 0) + h(1,4,3)$.

1.14 Let $\vec{r} = \vec{a}+k\vec{u}$, $\vec{r} = \vec{b}+h\vec{v}$, where $\vec{a}$, $\vec{b}$, $\vec{u}$, $\vec{v}$ are given vectors and h and k are parameters, represent two lines in 3-space. Show that if the lines intersect then $[\vec{a},\vec{u},\vec{v}] = [\vec{b},\vec{u},\vec{v}]$.

1.15 Find the orthogonal projection of the vector (3,4,5) into the plane $x-y-z = 0$. Answer. (5,2,3)

1.16 Give a geometric proof of Schwarz's inequality

$$(a_1b_1+a_2b_2+a_3b_3)^2 \leqq (a_1^{\,2}+a_2^{\,2}+a_3^{\,2})(b_1^{\,2}+b_2^{\,2}+b_3^{\,2})$$

with the equal sign holding only when $a_1/b_1 = a_2/b_2 = a_3/b_3$.

1.17 In Example 1.2 let the angle α vary with time so that $\frac{d\alpha}{dt} = \dot{\alpha}$ is not zero and each element of mass has a velocity component $s\dot{\alpha}\vec{j}$ in addition to the $\omega s \sin\alpha(-\vec{k})$ component considered in Example 1.2. Compute the moment of momentum of the rod with respect to 0.

1.18 Give three physical laws which involve a cross product.

1.19 A thin rod forming a conical pendulum is pinned at 0 to a support PO of length R which rotates with constant angular velocity ω. The rod has length L and mass M. Compute the moment of momentum of the rod with respect to point P, assuming that α is constant.

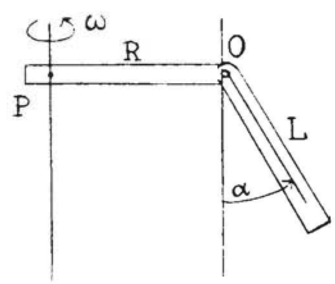


Figure 1.11

2. Derivatives and Partial Derivatives

Let the vector $\vec{u}$ in V_3 be a function of some continuous parameter t (say, position as a function of time). Then we may differentiate* $\vec{u}$ with respect to t simply by differentiating each component,

*We assume throught this section that our functions are smooth and so possess derivatives.

(2.1) $$\frac{d\vec{u}}{dt} = \left(\frac{du_1}{dt}, \frac{du_2}{dt}, \frac{du_3}{dt}\right).$$

Rule (2.1) leads immediately to the conclusion that the differential operator $\frac{d}{dt}$ is a linear operator on the vector space of vector-valued functions of t; that is,

(2.2) $$\frac{d}{dt}\ (a\vec{u} + b\vec{v}) = a\,\frac{d\vec{u}}{dt} + b\,\frac{d\vec{v}}{dt}\ ,$$

where $\vec{u},\vec{v}$ are vectors and a,b are constants. Moreover the familiar rule for differentiating a product holds for both the inner product and the cross product.

(2.3) $$\frac{d}{dt}\ (\vec{u}\cdot\vec{v}) = \frac{d\vec{u}}{dt}\cdot\vec{v} + \vec{u}\cdot\frac{d\vec{v}}{dt}\ ,$$

(2.4) $$\frac{d}{dt}\ (\vec{u}x\vec{v}) = \frac{d\vec{u}}{dt}\ x\vec{v} + \vec{u}x\ \frac{d\vec{v}}{dt}\ .$$

We will prove (2.4), using $\dot{x}$ for $\frac{dx}{dt}$ in the argument. It is sufficient to look at the first components on both sides of (2.4) since the other two components will be obtained by cyclic permutation of suffixes. On the left we have

$$\frac{d}{dt}\ (u_2v_3-u_3v_2) = \dot{u}_2v_3+u_2\dot{v}_3-\dot{u}_3v_2-u_3\dot{v}_2\ ,$$

while on the right we have

$$\dot{u}_2v_3-\dot{u}_3v_2+u_2\dot{v}_3-u_3\dot{v}_2\ .$$

These are evidently equal and so (2.4) is proved.

The reader should notice that $(1/h)(\vec{u}(t_0+h)-\vec{u}(t_0))$ is a vector function of h and its limit as $h \to 0$ is just the value of the vector function $\frac{d\vec{u}}{dt}$ at $t = t_0$. It is often convenient to think of a vector function without explicit reference to its components, especially when geometrical or physical ideas are in question. For example, let $\vec{r} = \vec{r}(t)$ describe a smooth curve in 3-space. Then, at any point on the curve, $\frac{d\vec{r}}{dt}$ is in the positive direction of the tangent to the curve and measures the velocity along the curve. Similarly, $\frac{d\vec{v}}{dt} = \vec{a}$, the acceleration. In this spirit the proof of (2.4) would proceed as follows:

$$\begin{aligned}\frac{d}{dt}(\vec{u}x\vec{v}) &= \lim_{\Delta t\to 0} \frac{(\vec{u}+\overrightarrow{\Delta u})x(\vec{v}+\overrightarrow{\Delta v}) - \vec{u}x\vec{v}}{\Delta t} \\ &= \lim_{\Delta t\to 0}(\vec{u}x\,\frac{\overrightarrow{\Delta v}}{\Delta t} + \frac{\overrightarrow{\Delta u}}{\Delta t}\,x\vec{v} + \overrightarrow{\Delta u}x\,\frac{\Delta\vec{v}}{\Delta t}) \\ &= \vec{u}x\,\frac{d\vec{v}}{dt} + \frac{d\vec{u}}{dt}\,x\vec{v}.\end{aligned}$$

However, in this 'proof', there are certain 'obvious' steps (e.g., $\lim_{\Delta t\to 0} \vec{u}x\,\frac{\overrightarrow{\Delta v}}{\Delta t} = \vec{u}x\,\frac{d\vec{v}}{dt}$), which really need mathematical justification .

<u>Example 2.1</u>. A point moves on a curve in a plane such that $r = 3t$, $\theta = 3t$ (in polar coordinates). Determine the velocity and acceleration when $t = 2\pi/3$.

The position vector $\vec{r}$ at time t is

$$\vec{r} = 3t\vec{e}_1$$

where $\vec{e}_1$ is a unit vector

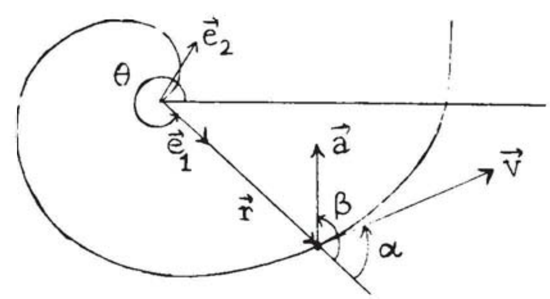


Figure 2.1

outward along $\vec{r}$. The velocity is

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{d}{dt}(3t\vec{e}_1)$$

or

$$\vec{v} = 3\vec{e}_1 + 3t\frac{d\vec{e}_1}{dt} = 3\vec{e}_1 + 9t\vec{e}_2,$$

where $\frac{d\vec{e}_1}{dt} = \dot{\theta}\vec{e}_2 = 3\vec{e}_2$, $\vec{e}_2$ being a unit vector perpendicular to $\vec{e}_1$.

When $t = 2\pi/3$, $|\vec{v}| = \sqrt{3^2 + (6\pi)^2}$ and $\tan\alpha = 6\pi/3 = 2\pi$. The acceleration is

$$\vec{a} = \frac{d\vec{v}}{dt} = 3\frac{d\vec{e}_1}{dt} + 9\vec{e}_2 + 9t\frac{d\vec{e}_2}{dt},$$

or

$$\vec{a} = 18\vec{e}_2 - 27t\vec{e}_1$$

because $\frac{d\vec{e}_2}{dt} = -\dot{\theta}\vec{e}_1 = -3\vec{e}_1$. When $t = 2\pi/3$, $|\vec{a}| = \sqrt{18^2+(18\pi)^2}$ and $\tan\beta = 18/(-18\pi) = -1/\pi$.

An important equation in dynamics is obtained by differentiating equation (1.13). Assuming that the mass m of the particle remains constant, we get

$$\frac{d\vec{H}}{dt} = m\frac{d}{dt}(\vec{r}\times\vec{v})$$

$$= m\left(\frac{d\vec{r}}{dt}\times\vec{v} + \vec{r}\times\frac{d\vec{v}}{dt}\right)$$

$$= m\,(\vec{v}\times\vec{v} + \vec{r}\times\frac{d\vec{v}}{dt})$$

$$= \vec{r}\times(m\,\frac{d\vec{v}}{dt}),$$

since $\vec{v}\times\vec{v} = 0$. But by Newton's Law, $m\,\frac{d\vec{v}}{dt}$ is equal to the force $\vec{F}$ acting on the particle, and so we have

(2.5) $$\frac{d\vec{H}}{dt} = \vec{r}\times\vec{F} = \vec{L}\ .$$

Example 2.2. Central-Force Motion.

Consider the motion of a particle on which the resultant force is always directed along the line between the particle and a fixed point 0. Such a central force $\vec{F}$ is collinear with $\vec{r}$, and so $\vec{L} = \vec{r}\times\vec{F} = 0$. From (2.5) it then follows that the moment of momentum $\vec{H}$ is constant. Since $\vec{H} = \vec{r}\times(m\vec{v})$, $\vec{r}$ and $\vec{v}$ must lie in the plane through 0 perpendicular to $\vec{H}$ (Figure 2.2). Thus, all central-force motions are planar and the moment of momentum is conserved. Using polar coordinates (r,θ) in the plane of motion, the velocity $\vec{v}$ has a component $\dot{r}$ outward along the radius and a component $r\dot{\theta}$ perpendicular to the radius. Thus, the magnitude of $\vec{r}\times m\vec{v}$ is

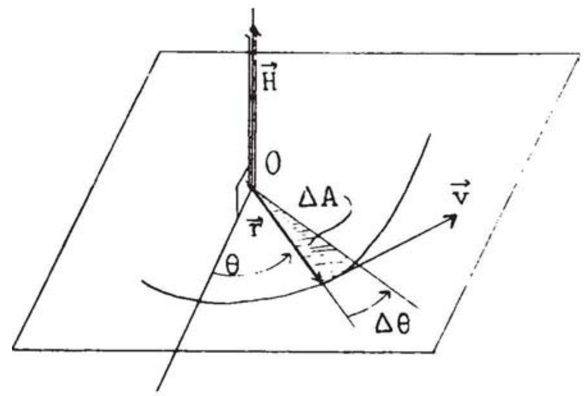


Figure 2.2

$$\|\vec{r}\times m\vec{v}\| = mr^2\dot{\theta} = \|\vec{H}\|, \qquad \text{a constant .}$$

The element of area dA swept out by position vector $\vec{r}$ during a small time increment Δt is given by

$$\Delta A = \frac{1}{2} r^2 \Delta\theta$$

Dividing by Δt and taking the limit as $\Delta t \to 0$ gives

$$\frac{dA}{dt} = \frac{1}{2} r^2\dot{\theta}.$$

But from above, $r^2\dot{\theta} = \frac{\|\vec{H}\|}{m}$ is a constant for central-force motion so that the areal velocity is constant for such motion. This is Kepler's second law for the case of a gravitational central force. (Compare this discussion with Problem 8.7 of Chapter 4.)

Example 2.3. Conical Pendulum.

In Example 1.2 the moment of momentum of the conical pendulum was found to be

$$\vec{H} = \frac{ML^2\omega \sin\alpha}{3} \vec{j}.$$

The time rate of change of $\vec{H}$ is

$$\frac{d\vec{H}}{dt} = \frac{ML^2\omega \sin\alpha}{3} \frac{d\vec{j}}{dt}$$

and

$$\frac{d\vec{j}}{dt} = \omega \cos\alpha \; (-\vec{k})$$

so that

$$\frac{d\vec{H}}{dt} = \frac{ML^2\omega^2\sin\alpha \cos\alpha}{3} (-\vec{k}).$$

The time rate of change of the moment of momentum equals the moment $\vec{L}$ of the external force with respect to point 0. Hence

$$\vec{L} = \frac{MgL\sin\alpha}{2}(-\vec{k}) = \frac{ML^2\omega^2\sin\alpha\cos\alpha}{3}(-\vec{k}),$$

and either $\alpha = 0$ or

$$\frac{3g}{2L} = \omega^2\cos\alpha \quad .$$

If $\omega^2 > 3g/2L$ there are two possible values for α, namely 0 and $\arccos(3g/2L\omega^2)$. If, on the other hand, $\omega^2 \leqq 3g/2L$ then only the trivial case $\alpha = 0$ is possible.

What has been said so far extends in an obvious way to partial derivatives. We will be content here to recall the basic definitions for ordinary real-valued functions and the reader will then have no difficulty in extending them to vector-valued functions analogous to (2.1).

Let $f(x,y,z)$ be a function of 3 real variables defined over some region S of 3-space and taking real values. (We use the symbols x,y,z, rather than x_1,x_2,x_3 for notational convenience -- and to emphasize the reader's familiarity with the topic!) Let (a,b,c) be a point of S. Then, by restricting the function f to the part of S lying along the line $y = b$, $z = c$, we obtain a function $g(x)$, defined by $g(x) = f(x,b,c)$. By definition, the value of the partial derivative $\frac{\partial f}{\partial x}$ at the point (a,b,c) is just $g'(a)$, where $g'(x)$ is the derivative of $g(x)$. The partial derivatives $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial z}$ are defined similarly. Thus, since partial derivatives are just ordinary derivatives (of functions obtained from the given function), all the usual properties of ordinary derivatives go over immediately to partial derivatives. So, in particular, the analogs of (2.2), (2.3), (2.4) hold for vector-valued functions of 3 real variables.

We should point out, however, that although the definition just given is, in fact, the precise definition of the partial derivative $\frac{\partial f}{\partial x}$, <u>in practice</u> one uses the simplified procedure usually described as 'differentiating with respect to x holding y and z fixed.' What we have done above is, essentially, to make the logical meaning of this procedure quite explicit. Thus to obtain $\frac{\partial f}{\partial x}$ when f is the function $e^{3x+2y} \cos z$, we regard y and z as constants and so $\frac{\partial f}{\partial x} = 3e^{3x+2y} \cos z$. The value of $\frac{\partial f}{\partial x}$ at the point (a,b,c) is then $3e^{3a+2b} \cos c$. We would not, in practice, construct the function $g = e^{3x+2b} \cos c$ and then differentiate it. However, as we have pointed out, the precise definition given immediately justifies many of the properties of partial derivatives in the light of their validity for ordinary derivatives. It also should serve to dissipate the feeling that partial derivatives involve more difficult concepts than ordinary derivatives.

However, there is one formula which has to be re-examined when we pass from ordinary derivatives to partial derivatives and this is the chain-rule for handling a change of variables.

Here the increase in the number of variables leads to a generalization of the formula.

Recall the chain rule for functions of a single variable. If $\varphi = \varphi(t)$ and $f = f(\varphi)$ then f may be regarded as a function of t, $f = f(\varphi(t))$ and

$$\frac{df}{dt} = \frac{df}{d\varphi}\frac{d\varphi}{dt}.$$

The generalization is as follows. Let f be a function of x,y,z; thus $f = f(x,y,z)$, and let each of x,y,z be functions of several variables t,u,v,...; we do not specify the number of variables since we will only be concerned with the partial derivatives with respect to t. Then f may be regarded as a function of the variables t,u,v,...,

$$f = f(x(t,u,v,\ldots), y(t,u,v,\ldots), z(t,u,v,\ldots)),$$

and the chain rule asserts

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}. \tag{2.6}$$

To avoid confusion it is sometimes better to use a different symbol for f as a function of t,u,v,... from the symbol used for f as a function x,y,z since they are certainly different <u>functions</u>. However, (2.6) is the natural form in which to remember the chain rule so we state it in this form, but with the warning. Of course, on the right hand side of (2.6), $\frac{\partial f}{\partial x}$ is first computed as a function of (x,y,z) and we then substitute for x,y,z as functions of t,u,v,... to get equality on the two sides of (2.6). We also emphasize that, in (2.6), $\frac{\partial f}{\partial t}$ is computed holding u,v,... fixed, while $\frac{\partial f}{\partial x}$, for example, is computed holding y,z fixed.

Let us take an example to clarify (2.6). Let $f(x,y,z) = x \sin z + y \cos z$, and let x,y,z be functions of θ, $x = a \cos\theta$, $y = a \sin\theta$, $z = 2\theta$. Then

$$\frac{\partial f}{\partial x} = \sin z,$$

$$\frac{\partial f}{\partial y} = \cos z,$$

$$\frac{\partial f}{\partial z} = x \cos z - y \sin z,$$

$$\frac{dx}{d\theta} = -a \sin\theta,$$

$$\frac{dy}{d\theta} = a \cos\theta,$$

$$\frac{dz}{d\theta} = 2.$$

If we express $f(x,y,z)$ as a function of θ we obtain the function $a \cos\theta \sin 2\theta + a \sin\theta \cos 2\theta = a \sin 3\theta$, whose derivative is $3a \cos 3\theta$. We may substitute into $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ to obtain functions of θ, and (2.6) then asserts that

$$(2.7) \quad 3a \cos 3\theta = -a \sin 2\theta \sin\theta + a \cos 2\theta \cos\theta + 2a(\cos\theta \cos 2\theta - \sin\theta \sin 2\theta) .$$

The reader will readily verify that (2.7) is, in fact, true, since $\cos 2\theta \cos\theta - \sin 2\theta \sin\theta = \cos 3\theta$.

A proof of (2.6) may be found, for example, in Thomas, Section 14-7.

Our discussion of partial derivatives has concerned itself with functions $f(x,y,z)$ of 3 real variables, but there is, of course, no reason from the analytical point of view to confine attention to the case of 3 variables; it is quite obvious, for example, how (2.6) extends to the case of n variables. However we have preferred to discuss the case of 3 variables since it has a special <u>geometrical</u> significance in view of our interest in this chapter in Euclidean space of 3 dimensions. Since we have supposed given a fixed coordinate system in E_3, a function of 3 variables may be identified with a function of a <u>vector</u> variable $\vec{r}$; we may write $f(\vec{r})$ for $f(x,y,z)$ if x,y,z are the components of $\vec{r}$. This point of view is very important to us; in particular if we consider functions of the variable $\vec{r}$, then this puts less emphasis on the coordinate axes. We are thus led to the natural notion of the partial derivative of f in a given direction in E_3, or the <u>directional</u> <u>derivative</u> of f in the given direction. Let the given direction be given by the unit vector $\vec{u}$. We consider, for fixed $\vec{r}$, the expression $f(\vec{r}+t\vec{u})$, where t is a real variable or parameter. As we saw in the preceding section, the vector $\vec{r}+t\vec{u}$ then describes the line through the point P with position vector $\vec{r}$ and in the direction of the vector $\vec{u}$. Then $f(\vec{r}+t\vec{u})$ is a function of the real variable t and we may obtain the derivative of this function at $t = 0$,

$$\left[\frac{d}{dt} f(\vec{r}+t\vec{u}) \right]_{t=0} . \tag{2.8}$$

The expression (2.8) then measures the rate of change of the function f in the direction $\vec{u}$ at the point P, and is called the directional derivative (in the direction $\vec{u}$).

Example 2.4. It is evident that, in principle, we can discuss directional derivatives for functions of any number of variables. In the present example we consider functions of x and y only, or, in other words, consider $\overrightarrow{OP} = \vec{r}$ as a position vector in a 2-dimensional Euclidean space E_2. Suppose that there is a hill rising above E_2, and let $h = f(\vec{r})$ be the elevation of the terrain above the point P. (Figure 2.3.) Then (2.8) measures the slope in the direction $\vec{u}$, i.e. the rate of increase of h as P moves in the direction given by $\vec{u}$. The minimum slope gives the "most downhill" direction, which is the direction a tiny ball placed at Q will roll under the action of gravity.

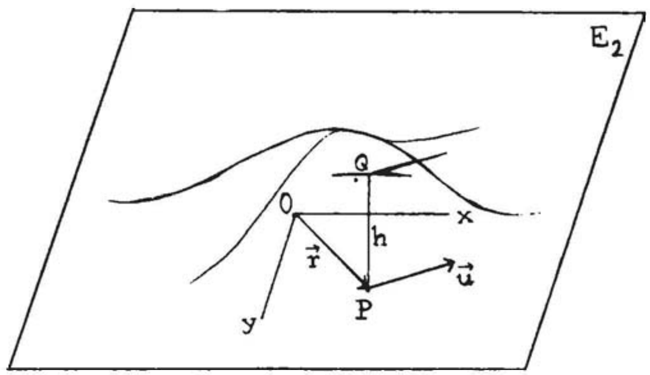


Figure 2.3

Let us write (u_1,u_2,u_3) for the components of $\vec{u}$ and (x,y,z) for the components of $\vec{r}$. Then

$$f(\vec{r}+t\vec{u}) = f(x+tu_1,\ y+tu_2,\ z+tu_3)$$

and (2.6) shows that

$$(2.9) \qquad \left[\frac{d}{dt} f(\vec{r}+t\vec{u})\right]_{t=0} = \frac{\partial f}{\partial x} u_1 + \frac{\partial f}{\partial y} u_2 + \frac{\partial f}{\partial z} u_3 ,$$

where the partial derivatives on the right are evaluated at (x,y,z). Notice that these partial derivatives $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial z}$ are independent of the direction $\vec{u}$.

The vector $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})$ plays a very important role in the vector differential calculus. It is called the gradient of f and written gradf or ∇f. We will have much more to say about it in the next section, but meanwhile we record (2.9) in the form of a theorem.

Theorem 2.1. The directional derivative of the function f in the direction $\vec{u}$ is the inner product

$$\vec{u} \cdot \text{grad} f, \quad \text{or} \quad \vec{u} \cdot \nabla f.$$

As an application of this theorem consider a curve C in E_3 given by $\vec{r} = \vec{r}(t)$. If P is a point on this curve, we may ask for the directional derivative of f at P in the direction of the tangent to the curve. Now if s is arc length measured along the curve, then the coordinates x,y, and z of P are functions of the one variable s, and the chain-rule (2.6) becomes

$$\frac{df}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} + \frac{\partial f}{\partial z}\frac{dz}{ds},$$

or

$$\frac{df}{ds} = \frac{d\vec{r}}{ds} \cdot \text{grad } f. \tag{2.10}$$

Since $\frac{d\vec{r}}{ds}$ is the unit tangent vector to C (see Thomas, Section 13-8), the directional derivative of f in the direction of the tangent is the same as the derivative of f considered as a function of the arc length.

If we wish to avoid the parameter s in favor of t we need merely note that since

$$\frac{d\vec{r}}{dt} = \frac{d\vec{r}}{ds}\frac{ds}{dt}, \quad \text{and} \quad \left\|\frac{d\vec{r}}{ds}\right\| = 1,$$

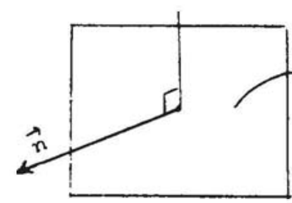
we have $\left\|\frac{d\vec{r}}{dt}\right\| = \frac{ds}{dt}$, (assuming s measured so as to increase with t). Then (2.10) can be written as

$$\text{(2.11)} \qquad \frac{df}{ds} = \left(\frac{d\vec{r}}{dt} \Big/ \left\|\frac{d\vec{r}}{dt}\right\|\right) \cdot \text{grad} f \; .$$

Example 2.5. Pressure in an Inviscous Fluid

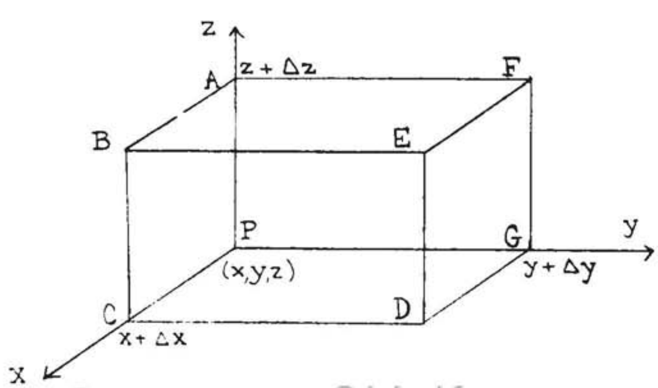
Consider an inviscous fluid. By inviscous we mean that the fluid cannot support a shear stress. In other words if we consider a small element of surface dS in the fluid then the force exerted by the fluid on one side of dS on the fluid on the other side is normal to the surface dS.

Figure 2.4



The pressure at any point (x,y,z) in the fluid is a scalar p(x,y,z). Consider the box shown in Figure 2.5.

Figure 2.5



The force acting on the face PABC is $\int_{z}^{z+\Delta z}\int_{x}^{x+\Delta x} p(\xi,y,\zeta)d\xi d\zeta\ \vec{j}$.

Similarly the force on the face DEFG is

$$-\int_{z}^{z+\Delta z}\int_{x}^{x+\Delta x} p(\xi,y+\Delta y,\zeta)d\xi d\zeta\ \vec{j}\ .$$

1

2 3

The net force in the $\vec{j}$ direction is then

$$-\int_{z}^{z+\Delta z}\int_{x}^{x+\Delta x} [p(\xi,y+\Delta y,\zeta)-p(\xi,y,\zeta)]d\xi d\eta\ \vec{j}\ .$$

which by the mean value theorem for integrals[†] is

$$-\ [p(\xi^*,y+\Delta y,\zeta^*)-p(\xi^*,y,\zeta^*)](\Delta x)(\Delta z)\vec{j}$$

where ξ^* is between x and $x+\Delta x$ and ζ^* is between z and $z+\Delta z$.

The force per unit volume in the $\vec{j}$ direction is

$$-\ \frac{[p(\xi^*,y+\Delta y,\zeta^*)-p(\xi^*,y,\zeta^*)]}{\Delta y}\ \vec{j}\ .$$

If we now let $\Delta x \to 0$, $\Delta y \to 0$, $\Delta z \to 0$, the force per unit volume, in the $\vec{j}$ direction, at the point (x,y,z) is given by

[†]The mean value theorem for integrals says that the integral of a continuous function over a bounded, connected region is equal to the area of the region times the value of the function at some point in the region. For a more general statement and a proof see T. M. Apostol, Mathematical Analysis, Addison-Wesley Publishing Co., Reading, Mass., 1957, page 269.

$-\frac{\partial p}{\partial y}\vec{j}$. Similarly the forces per unit volume in the $\vec{i}$ and $\vec{k}$ directions at the point (x,y,z) are given by $-\frac{\partial p}{\partial x}\vec{i}$ and $-\frac{\partial p}{\partial z}\vec{k}$. Thus the total force per unit volume is given by

$$-\left(\frac{\partial p}{\partial x}\vec{i} + \frac{\partial p}{\partial y}\vec{j} + \frac{\partial p}{\partial z}\vec{k}\right) = -\nabla p ,$$

i.e., the force per unit volume at a point is the negative of the pressure gradient at that point.

Example 2.6. Heat Flow

Consider a solid body. The temperature at a point (x,y,z) in the body is given by u(x,y,z). If we consider a surface dS in the body, with unit normal $\vec{n}$ then the rate at which heat flows across the surface is known to be proportional to the derivative of the temperature in the direction $\vec{n}$ and proportional to the area dS. It flows of course from the higher to the lower temperature. It follows that we can regard the flow of heat as a vector $\vec{q} = -\kappa\nabla u$, and the flow across the surface dS is given by $\vec{q}\cdot(\vec{n}dS) = -\kappa\nabla u\cdot(\vec{n}dS)$ where κ is called the coefficient of thermal conductivity. This fact was used in the heat conduction problems of Chapter 1.

Problems

2.1 Use the chain-rule to solve the following problems.

(a) Find $\frac{\partial f}{\partial t}$ if $f = x^4 + yz^3$, $x = s^2-t^2$, $y = 2st$, $z = s^2+t^2$.

(b) Find $\frac{d\theta}{ds}$ if θ is arctan $\frac{y}{x}$, $x = \sqrt{1+s^2}$, $y = \sqrt{1-s^2}$.

(c) Find $\frac{dz}{dx}$ if $z = x \sin(x-y)$, $y = \arcsin x$.

[Consider $z = u \sin(u-y)$ where $u = x$, $y = \arcsin x$.]

(d) Find $\frac{dz}{dx}$ if $z = x \sin(x-y)$, $x^3+xy+y^3 = 1$.

[Consider the last equation as defining y as an implicit function of x.]

2.2 If $\vec{u},\vec{v},\vec{w}$ are vector functions of t show that

$$\frac{d}{dt}[\vec{u},\vec{v},\vec{w}] = [\frac{d\vec{u}}{dt},\vec{v},\vec{w}] + [\vec{u},\frac{d\vec{v}}{dt},\vec{w}] + [\vec{u},\vec{v},\frac{d\vec{w}}{dt}] .$$

2.3 Let pressure, temperature and volume be related by the equation $f(p,t,v) = 0$, in such a way that each variable can be expressed as a differentiable function of the other two. Let $(\frac{\partial p}{\partial t})_v$ stand for the partial derivative of pressure with respect to temperature, volume being held fixed, and let $(\frac{\partial t}{\partial v})_p$, $(\frac{\partial v}{\partial p})_t$, be defined similarly.

(a) Prove that $(\frac{\partial p}{\partial t})_v = -(\frac{\partial f}{\partial t})/(\frac{\partial f}{\partial p})$, with similar expressions for $(\frac{\partial t}{\partial v})_p$ and $(\frac{\partial v}{\partial p})_t$.

(b) Prove that $(\frac{\partial p}{\partial t})_v (\frac{\partial t}{\partial v})_p (\frac{\partial v}{\partial p})_t = -1$.

Check this equation when $f = pv - kt$, k being a constant.

2.4 Find the directional derivative of the function $z-x^2-y^2$ in the direction of the vector $(1,2,2)$ at the point $(2,-1,1)$. Find also the gradient of the function at this point.

2.5 Find the directional derivative of the function $\frac{1}{x^2+y^2+z^2}$ in the direction of the vector $(2,3,-6)$ at the point $(1,4,1)$. Also find the gradient of the function at this point.
Answer. $-4/567$, $-(i + 4j + k)/162$.

2.6 Find the derivative of the function $x^2+y^2+z^2$ along the curves

(a) $x = a \cos t$, $y = a \sin t$, $z = bt$; Ans. $2b^2t/\sqrt{a^2+b^2}$.

(b) $x = a \cos t$, $y = a \sin t$, $z = b \cosh \frac{at}{b}$;

(c) $x = \log \cos t$, $y = \log \sin t$, $z = \sqrt{2t}$.

2.7 Suppose the "hill" in Example 2.4 is given by $h = f(x,y)$, then $\nabla f = \frac{\partial f}{\partial x}\vec{i} + \frac{\partial f}{\partial y}\vec{j}$ gives the direction of a freely moving particle at each point. The "streamlines" of the flow are tangent to this direction. Thus if $y = y(x)$ represents a streamline we have $\frac{dy}{dx} = \frac{\partial f/\partial y}{\partial f/\partial x}$

(a) Find the streamlines if the hill is defined by

$$h = \frac{1}{x^2+4y^2}. \quad (\text{Ans. } y = cx^4.)$$

If $y = y(x)$ defines a level curve, or contour line, then $z = f(x,y) = C$ so that $\frac{\partial f}{\partial x} + \frac{\partial f}{\partial x}\frac{dy}{dx} = 0$ and $\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y}$.

(b) From this show that the streamlines are orthogonal to the contour lines. (The streamlines are called the "orthogonal trajectories" of the contour lines.)

(c) Find the contour lines when $h = \dfrac{1}{x^2+4y^2}$. Ans. $x^2+4y^2 =$

(d) Sketch the streamlines found in (a) and the contour lines found in (c) on the same graph.

2.8 A point P moves on curve C such that $r = e^t$, $\theta = \sqrt{3}\,(1-e^{-t})$ where r, θ are the coordinates shown in Figure 2.6. Starting with the position vector

$$\vec{r} = e^t \vec{e}_1$$

determine expressions for the velocity and acceleration. Evaluate these when $t = 0$ showing the vectors on a diagram.

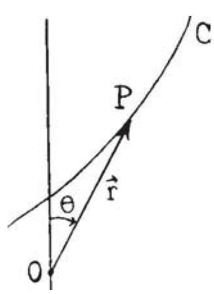


Figure 2.6

2.9 A missile M pursuing an airplane A is guided so that it is always headed towards the airplane, and so that its speed v ($=\|\vec{v}\|$) is proportional to its distance D from the airplane ($v = KD$). (See Figure 2.7)

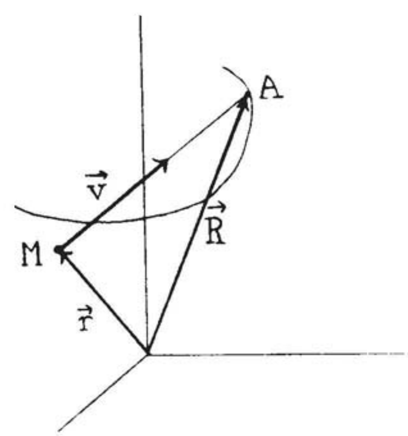


Figure 2.7

Suppose that the path of the airplane is the helix

$$\vec{R} = (\cos t)\vec{i} + (\sin t)\vec{j} + t\vec{k},$$

and that the missile starts at the origin ($\vec{r} = 0$ when $t = 0$).

(a) Find the differential equation describing the path $\vec{r}(t)$ of the missile.

(b) Solve the equation to find the path.

(c) Find the distance D as a function of t.

(d) Does the missile catch the plane?

Answers. (a) $\frac{d\vec{r}}{dt} = K(\vec{R}-\vec{r})$.

(b) $$\vec{r} = \frac{K}{K^2+1}(K\cos t + \sin t - Ke^{-Kt})\vec{i}$$

$$+ \frac{K}{K^2+1}(K\sin t - \cos t + e^{-Kt})\vec{j} + \frac{1}{K}(Kt - 1 + e^{-Kt})\vec{k}.$$

(c) $$D(t)^2 = \frac{1}{K^2+1}(1 - 2Ke^{-Kt}\sin t + K^2e^{-2Kt})$$

$$+ \frac{1}{K^2}(1-e^{-Kt})^2 .$$

2.10 Let $\vec{\omega} = \omega_1\vec{i} + \omega_2\vec{j} + \omega_3\vec{k}$ be a fixed vector and let $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$ as usual.

(a) Show that the relation $\vec{r}' = \vec{\omega}\times\vec{r}$ is expressible in the matrix form

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \text{ where } A = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

(b) Show that the eigenvalues of A are 0, $\pm i\omega$, where $\omega = \|\vec{\omega}\|$.

(c) Show that $\vec{\omega}$ is an eigenvector of A associated with the eigenvalue 0.

(d) For the special case $\vec{\omega} = 3\vec{i} + 4\vec{j}$, find an eigenvector associated with the eigenvalue $i\omega$.

(e) Use the above results to get a solution of the differential equation $\frac{d\vec{r}}{dt} = \vec{\omega} x \vec{r}$, where $\vec{\omega}$ has the value given in (d).

2.11 Let $\vec{r} = (\sqrt{2}\cos 3t,\ 1 + \sin 3t,\ 1 - \sin 3t)$ be the radius vector from the origin to a moving point P.

(a) Show that $\vec{r}\cdot\vec{r}$ = constant. What geometrical property of the locus of the point P does this imply?

(b) Compute the velocity $\vec{v} = \frac{d\vec{r}}{dt}$ of P. Show by direct computation that $\vec{v}$ is orthogonal to $\vec{r}$.

(c) Show that $\vec{v}$ is orthogonal to $\vec{r}$ by differentiating the equation $\vec{r}\cdot\vec{r} = c$ with respect to t.

(d) Prove that the locus of P is a circle.

3. Scalar and Vector Fields and the Differential Operator ∇.

A function φ defined over the whole of 3-space E_3, or over a region of E_3, is called a scalar field if the values taken by the function are real numbers.* We always insist that the function φ be continuous in talking of a field; in fact, we will here assume further whatever differentiability conditions

*One can, of course, discuss complex scalar and vector fields, but these will not enter into our considerations.

are needed to make sense of our statements and definitions.

We now introduce the important notion of a <u>vector field</u> over E_3, or over a region of E_3. This is again a function $\vec{u}$ defined over E_3 (or a region of E_3) but $\vec{u}$ is now a vector-valued function; that is, the values of $\vec{u}$ are to be vectors in E_3, or, equivalently, V_3. We recalled, in the previous section, how to extend the notions of the partial differential calculus from real-valued functions to vector-valued functions, that is, from scalar fields to vector fields.

At the end of the last section we met the gradient of a function. We see now that we may regard the gradient as a sort of vector differential operator, transforming a scalar field φ into a vector field gradφ according to the formula

$$\nabla\varphi = \text{grad}\varphi = (\frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}, \frac{\partial\varphi}{\partial z}). \tag{3.1}$$

If we wish to set in prominence the differential operator itself, we may even write

$$\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}) . \tag{3.2}$$

We now use Theorem 2.1 to give a geometrical characterization of the gradient. Let $\vec{u}$ be a variable unit vector and let $\vec{v}$ be a fixed non-zero vector. Then

$$\vec{v}\cdot\vec{u} = \|\vec{v}\| \cos\alpha ,$$

where α is the angle between the directions of $\vec{u}$ and $\vec{v}$. Thus $\vec{u}\cdot\vec{v}$ attains its maximum value, namely $\|\vec{v}\|$, when $\alpha = 0$, that is, when $u = \frac{\vec{v}}{\|\vec{v}\|}$. Now the directional derivative of φ in the direction $\vec{u}$ is, by Theorem 2.1, just grad$\varphi\cdot\vec{u}$. Thus the directional derivative is maximum when $\vec{u}$ is in the direction of gradφ and its value is then |gradφ|. We sum this up in the following theorem.

Theorem 3.1. Let φ be a scalar field. Then, at each point P, the direction of the gradient vector field gradφ is that along which the rate of change of φ is a maximum, and the magnitude of the vector field, |gradφ|, is that maximum rate of change.

The reader should consider Examples 2.4, 2.5 and 2.6 given in the previous section in the light of this theorem. He will note that the direction in which the ball will roll, the force on a particle in the fluid and direction of heat flow are each determined by the gradient vector field (of the appropriate function; the negative of the height, the pressure, or the temperature).

Example 3.1. Consider a scalar field which depends only on the distance of the point from the origin. This is often described as the condition of spherical symmetry and arises frequently in practice. Thus we are concerned with a function $\varphi(r)$ where $r = |\vec{r}|$ is the length of OP. If we compute the gradient we obtain

$$\text{grad}\varphi = \left(\frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}, \frac{\partial\varphi}{\partial z}\right) = \left(\frac{d\varphi}{dr}\frac{\partial r}{\partial x}, \frac{d\varphi}{dr}\frac{\partial r}{\partial y}, \frac{d\varphi}{dr}\frac{\partial r}{\partial z}\right) .$$

Now $r^2 = x^2+y^2+z^2$, so that $r\frac{\partial r}{\partial x} = x$, $r\frac{\partial r}{\partial y} = y$, $r\frac{\partial r}{\partial z} = z$. Thus

$$\operatorname{grad}\varphi = \frac{1}{r}\frac{d\varphi}{dr}(x,y,z) = \frac{1}{r}\frac{d\varphi}{dr}\vec{r} .$$

It follows that $\operatorname{grad}\varphi$ is a vector in the direction of the radius vector $\vec{r}$ and of magnitude $\frac{d\varphi}{dr}$. That is to say (and this should seem intuitively reasonable to the reader) the greatest rate of change of φ is along the radius vector and the rate of change along the radius vector is $\frac{d\varphi}{dr}$.

A particularly interesting property of the gradient field is the following. The equation $\varphi(\vec{r}) = c$ determines a surface in E_3, called an equipotential or <u>level</u> surface of the scalar field φ. For a given constant c, let S(c) be the corresponding level surface and let P_0 be a point of S(c). If C, given by $\vec{r} = \vec{r}(t)$, is any curve in the surface S(c) passing through P_0, then, of course, φ is constant along C so $\frac{d\varphi}{dt} = 0$. Since

$$\frac{d\varphi}{dt} = \frac{\partial\varphi}{\partial x}\frac{dx}{dt} + \frac{\partial\varphi}{\partial y}\frac{dy}{dt} + \frac{\partial\varphi}{\partial z}\frac{dz}{dt} = \frac{d\vec{r}}{dt}\cdot\operatorname{grad}\varphi ,$$

we have

$$\frac{d\vec{r}}{dt}\cdot\operatorname{grad}\varphi = 0 . \tag{3.3}$$

Equation (3.3) says that $\operatorname{grad}\varphi$ is perpendicular to the direction of C at P_0; but as C was <u>any</u> curve in the level surface S(c) at P_0, this implies (by definition of the normal) that $\operatorname{grad}\varphi$ is in the direction of the normal to the surface S(c) at P_0. The reader should consider this fact in the light of the characterization we gave above of the direction of $\operatorname{grad}\varphi$ as that direction in which φ is changing fastest.

We are now able to use the operator ∇ to obtain the tangent plane and normal line to a given surface. Let the equation of the surface be $f(x,y,z) = 0$. Then the direction of the normal at (x_0,y_0,z_0) is given by the vector grad $f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})$ evaluated at (x_0,y_0,z_0). Thus the vector equation of the normal to the surface at $\vec{r}_0 = (x_0,y_0,z_0)$ is, in terms of a parameter t,

$$\vec{r} = \vec{r}_0 + t(\text{grad} f)_0. \tag{3.4}$$

The tangent plane to the surface at (x_0,y_0,z_0) is just the set of all points (x,y,z) such that the line joining (x_0,y_0,z_0) to (x,y,z) is perpendicular to the normal. Thus the equation of the tangent plane to the surface at $\vec{r}_0$ is

$$(\vec{r}-\vec{r}_0)\cdot(\text{grad} f)_0 = 0 . \tag{3.5}$$

Notice that (3.4) is an equation between vectors and involves the parameter t, while (3.5) is a scalar equation.

It may, of course, happen that the equation of a surface is not given in the form $f(x,y,z) = 0$. One common possibility is that the equation is given in the form $z = g(x,y)$. However, this is easily reduced to the standard form by setting $f(x,y,z) = g(x,y)-z$. Then

$$\text{grad} f = (\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, -1)$$

so that the equations of the normal take the coordinate form

$$\frac{x-x_0}{(\frac{\partial g}{\partial x})_0} = \frac{y-y_0}{(\frac{\partial g}{\partial y})_0} = -(z-z_0) \tag{3.6}$$

and the equation of the tangent plane takes the coordinate form

(3.7) $$z-z_0 = \left(\frac{\partial g}{\partial x}\right)_0 (x-x_0) + \left(\frac{\partial g}{\partial y}\right)_0 (y-y_0),$$

the partial derivatives being evaluated at (x_0, y_0). Note the analogy to the equation $y-y_0 = \left(\frac{dy}{dx}\right)_0 (x-x_0)$ of the tangent line to a curve in a plane.

A second way to represent a surface is parametrically. That is, we have two parameters, say λ, μ in the surface S and points of the surface are given by

$$x = x(\lambda,\mu), \quad y = y(\lambda,\mu), \quad z = z(\lambda,\mu),$$

or $$\vec{r} = \vec{r}(\lambda,\mu).$$

For example the points on the sphere $x^2+y^2+z^2 = a^2$ may be represented parametrically by $x = a \cos\theta \sin\varphi$, $y = a \sin\theta \sin\varphi$, $z = a \cos\varphi$. (Here the angles θ and φ of spherical coordinates play the roles of λ and μ.) When a surface is given parametrically we may obtain the tangent plane and normal to the surface at a given point P_0 on the surface as follows. Consider the curve $\mu = c$ on the surface passing through P_0. This may also be described as the curve $\vec{r} = \vec{r}(\lambda, c)$ and so its direction is $\frac{\partial \vec{r}}{\partial \lambda}$. Thus the normal to the surface is perpendicular to the direction $\frac{\partial \vec{r}}{\partial \lambda}$. Similarly it is perpendicular to the direction $\frac{\partial r}{\partial \mu}$ and it is therefore in the direction $\frac{\partial \vec{r}}{\partial \lambda} \times \frac{\partial \vec{r}}{\partial \mu}$. The equation of the normal in terms of a parameter h is therefore

$$\vec{r} = \vec{r}_0 + h\left(\frac{\partial\vec{r}}{\partial\lambda} \times \frac{\partial\vec{r}}{\partial\mu}\right)_0 , \tag{3.8}$$

the subscript indicating that the quantities are evaluated at P_0. The tangent plane is just the plane through P_0 containing the directions $\frac{\partial\vec{r}}{\partial\lambda}$ and $\frac{\partial\vec{r}}{\partial\mu}$ and thus is represented by the equation

$$\left[\vec{r}-\vec{r}_0,\ \frac{\partial\vec{r}}{\partial\lambda},\frac{\partial\vec{r}}{\partial\mu}\right] = 0, \tag{3.9}$$

where the square brackets indicate the triple scalar product.

Let us take the example of the sphere $x^2+y^2+z^2 = a^2$ and calculate the tangent plane at the point (x_0,y_0,z_0) corresponding to the values θ_0, φ_0 of the parameters. (See Figure 3.1.)

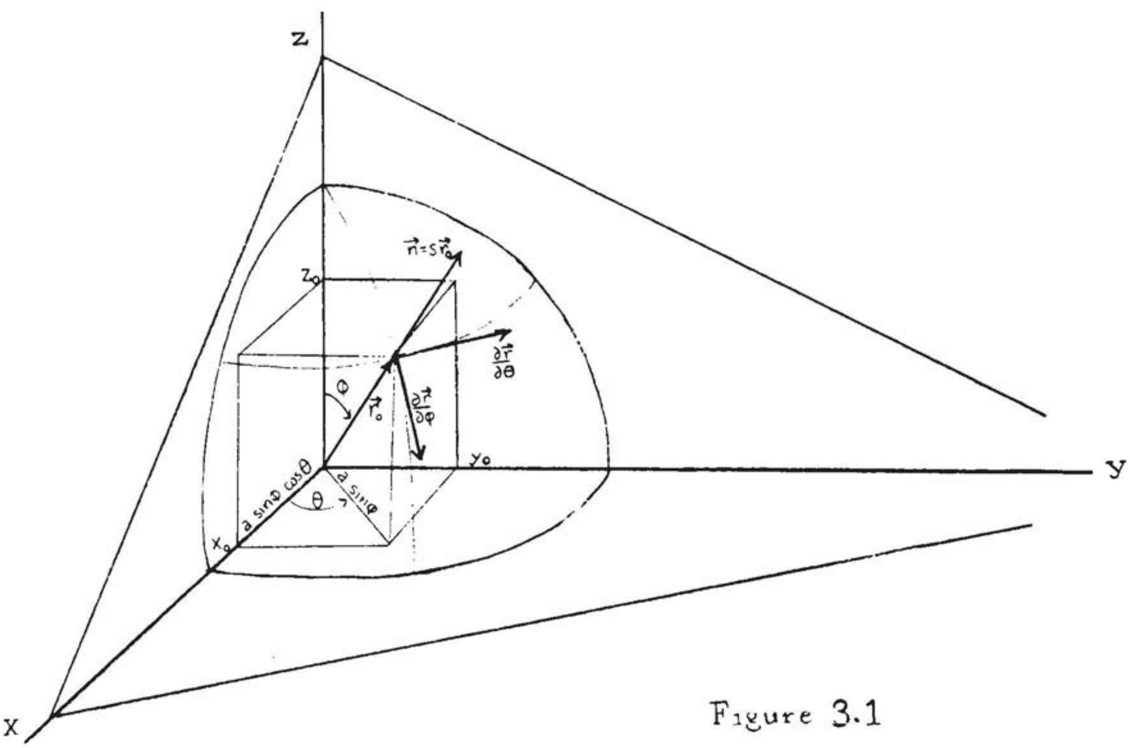


Figure 3.1

We compute:

$$\frac{\partial\vec{r}}{\partial\theta} = (-a\ \sin\theta\ \sin\varphi,\ a\ \cos\theta\ \sin\varphi,\ 0),$$

$$\frac{\partial\vec{r}}{\partial\varphi} = (a\ \cos\theta\ \cos\varphi,\ a\ \sin\theta\ \cos\varphi,\ -\ a\ \sin\varphi);$$

so that the normal vector is given by the cross product

$$\begin{aligned}\frac{\partial\vec{r}}{\partial\theta} \times \frac{\partial\vec{r}}{\partial\varphi} &= a^2(-\cos\theta\ \sin^2\varphi,\ -\ \sin\theta\ \sin^2\varphi,\ -\ \sin\varphi\ \cos\varphi)\\ &= -a^2\ \sin\varphi\ (\cos\theta\ \sin\varphi,\ \sin\theta\ \sin\varphi,\ \cos\varphi)\\ &= -(a\ \sin\varphi)\vec{r}.\end{aligned}$$

Thus the normal vector at any point P_0 is collinear with the radius vector $\vec{r}_0$ to that point - a well known property of the sphere. It follows that the equation of the tangent plane is $(\vec{r}-\vec{r}_0)\cdot\vec{r}_0 = 0$ or

$$\vec{r}\cdot\vec{r}_0 = \vec{r}_0\cdot\vec{r}_0 = \|\vec{r}_0\|^2 = a^2\ ;$$

so finally ,

$$\vec{r}\cdot\vec{r}_0 = a^2\ .$$

We revert now to the view contained in the formula (3.2)

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)\ .$$

According to this view, ∇ is to be thought of as a differential operator which, when applied to the scalar field φ, produces the gradient vector field grad$\varphi = \nabla\varphi$. Thus ∇ is a transformation from scalar fields to vector fields. We can regard ∇ as a sort of 'vector' whose 'components' are given by (3.2); $\nabla\varphi$ is then highly analogous to the usual multiplication of a vector and a scalar. The analogy is reinforced by the observation that ∇ is linear in that

$$\nabla(a\varphi+b\psi) = a\nabla\varphi + b\nabla\psi$$

for any constants a and b. On the other hand we must not press the analogy too far since certainly $\nabla(\varphi\psi) \neq (\nabla\varphi)\psi$. If, however, we do allow ourselves to think of ∇ as a vector, then it is natural to try out the usual algebraic operations on it, such as inner product and cross product. This may seem fanciful, but it is often a good procedure in mathematics to let the symbols take charge -- temporarily! -- and see whither they lead us. Let us do this -- we will in fact find ourselves led into notions of great physical and geometrical significance.

We are then looking at ∇ as a vector differential operator and we are naturally led to consider the expressions $\nabla\cdot\vec{u}$, $\nabla x\vec{u}$, where $\vec{u}$ is a vector field. The former is a scalar field given by

$$(3.10) \qquad \nabla\cdot\vec{u} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z},$$

where

$$\vec{u} = (u_1,u_2,u_3) \ ;$$

and the latter is a vector field given in components by

$$\text{(3.11)} \qquad \nabla x\vec{u} = (\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}, \frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}, \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}) ,$$

for the formulas (3.10), (3.11) are simply obtained by combining the 'vector' ∇ with the vector $\vec{u}$ by the usual rules for forming the inner and the cross product. We call $\nabla\cdot\vec{u}$ the <u>divergence</u> of $\vec{u}$ and sometimes write it div $\vec{u}$; we call $\nabla x\vec{u}$ the <u>curl</u> or <u>rotation</u> of $\vec{u}$ and sometimes write it curl $\vec{u}$.

Of course, the divergence and curl did not arise historically in the way we have described. Certain physical problems (particularly in hydrodynamics) gave rise to mathematical formulations which, in terms of a given Cartesian coordinate system, produced precisely the analytical expressions appearing on the right hand side of (3.10) and (3.11). It is not possible at the present stage to give a complete description of the physical significance of the divergence and the curl, but it should be worthwhile to sketch in some indications of their significance. We will begin with a discussion of the divergence; but first we make a remark, namely, that there is need to establish such a physical or geometrical significance even from the purely mathematical point of view. For, as matters stand, (3.10) and (3.11) attach meanings to $\nabla\cdot\vec{u}$ and $\nabla x\vec{u}$ <u>only when the coordinate system has been fixed</u>. We are naturally interested to know whether the quantities we get by applying the formulas on the right hand side of (3.10) and (3.11) correspond to something in the underlying Euclidean space. (For example, the formula

$\sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$ gives the distance between the points P, P' and thus has the same value no matter how we coordinatize E_3.) We will take up this question again in Section 5; at present we are content to have pointed out that there is a mathematical problem here as well as the obvious problem of attaching a physical significance to divergence and curl.*

Example 3.2. We consider first the flow of a fluid. At each point (x,y,z) and time t the velocity of the fluid is given by the vector

$$\vec{v} = v_1(x,y,z,t)\vec{i} + v_2(x,y,z,t)\vec{j} + v_3(x,y,z,t)\vec{k} ,$$

and the density is a scalar

$$\rho = \rho(x,y,z,t) .$$

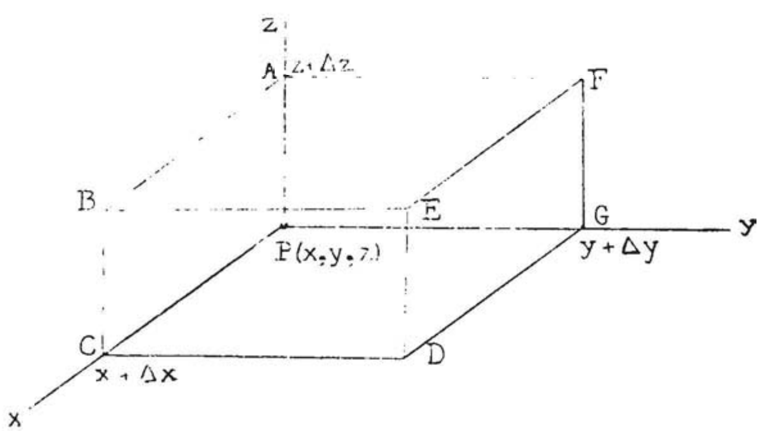


Figure 3.2

In a time interval $(t,t + \Delta t)$ the mass of fluid which flows into the box through the face PABC is

*The reader should note that we have already provided a purely geometrical interpretation of grad.

$$\int_{t}^{t+\Delta t}\int_{z}^{z+\Delta z}\int_{x}^{x+\Delta x} v_2(\xi,y,\zeta,\tau)\rho(\xi,y,\zeta,\tau)\,d\xi\,d\zeta\,d\tau$$

while that flowing out through the face DEFG is

$$\int_{t}^{t+\Delta t}\int_{z}^{z+\Delta z}\int_{x}^{x+\Delta x} v_2(\xi,y+\Delta y,\zeta,\tau)\rho(\xi,y+\Delta y,\zeta,\tau)\,d\xi\,d\zeta\,d\tau .$$

Thus, as far as these two faces are concerned we have

$$\text{Outflow-Inflow} = \int_{t}^{t+\Delta t}\int_{z}^{z+\Delta z}\int_{x}^{x+\Delta x} [F_2(\xi,y+\Delta y,\zeta,\tau)-F_2(\xi,y,\zeta,\tau)]\,d\xi\,d\zeta\,d\tau$$

where, for brevity we temporarily let $F_2 = \rho v_2$. By the mean value theorem for integrals we have, for these two faces,

$$\text{Outflow-Inflow} = [F_2(\xi^*,y+\Delta y,\zeta^*,\tau^*)-F_2(\xi^*,y,\zeta^*,\tau^*)](\Delta x)(\Delta z)(\Delta t)$$

where ξ^* is between x and $x+\Delta x$, ζ^* is between z and $z+\Delta z$, and τ^* is between t and $t+\Delta t$. Applying the same analysis to the other two pairs of opposite faces we find

(3.12) Total Outflow-Total Inflow =

$$[F_2(\xi^*,y+\Delta y,\zeta^*,\tau^*) - F_2(\xi^*,y,\zeta^*,\tau^*)](\Delta x)(\Delta z)(\Delta t) +$$

$$[F_1(x+\Delta x,\eta',\zeta',\tau') - F_1(x,\eta',\zeta',\tau')](\Delta y)(\Delta z)(\Delta t) +$$

$$[F_3(\xi'',\eta'',z+\Delta z,\tau'') - F_3(\xi'',\eta'',z,\tau'')](\Delta x)(\Delta y)(\Delta t)$$

where $F_1 = \rho v_1, F_3 = \rho v_3$, so that $\rho\vec{v} = (F_1,F_2,F_3)$.

Assuming that there is no source or sink inside the box, then from the conservation of mass we see that the [Total Outflow-Total Inflow] is equal to the negative of the change of mass of fluid in the box,

$$-\left[\int_z^{z+\Delta z}\int_y^{y+\Delta y}\int_x^{x+\Delta x} [\rho(\xi,\eta,\zeta,t+\Delta t)-\rho(\xi,\eta,\zeta,t)]d\xi d\eta d\zeta\,]\right].$$

Using the mean value theorem for integrals again we get

(3.13) Total Inflow-Total Outflow =

$$[\rho(\xi^{**},\eta^{**},\zeta^{**},t+\Delta t)-\rho(\xi^{**},\eta^{**},\zeta^{**},t)](\Delta x)(\Delta y)(\Delta z).$$

Equating the right hand sides of (3.12) and (3.13) and dividing by $(\Delta x)(\Delta y)(\Delta z)(\Delta t)$ we obtain

$$(3.14)\quad \frac{F_1(x+\Delta x,\eta',\zeta',\tau')-F_1(x,\eta',\zeta',\tau')}{\Delta x}$$

$$+\ \frac{F_2(\xi^*,y+\Delta y,\zeta^*,\tau^*)-F_2(\xi^*,y,\zeta^*,\tau^*)}{\Delta y}$$

$$+\ \frac{F_3(\xi'',\eta'',z+\Delta z,\tau'')-F_3(\xi'',\eta'',z,\tau'')}{\Delta z}$$

$$= -\left[\frac{\rho(\xi^{**},\eta^{**},\zeta^{**},t+\Delta t)-\rho(\xi^{**},\eta^{**},\zeta^{**},t)}{\Delta t}\right] .$$

As Δx, Δy, Δz, Δt tend to zero we obtain from (3.14)

$$\frac{\partial F_1}{\partial x}+\frac{\partial F_2}{\partial y}+\frac{\partial F_3}{\partial z} = -\frac{\partial\rho}{\partial t},\quad \text{or}\quad \nabla\cdot(F_1,F_2,F_3) = -\frac{\partial\rho}{\partial t} .$$

Since $(F_1,F_2,F_3) = \rho\vec{v}$, we obtain the Equation of Continuity:

(3.15) $$\nabla\cdot(\rho\vec{v}) = -\frac{\partial\rho}{\partial t}\ .$$

If the fluid is incompressible, so that ρ is constant this equation becomes simply

(3.16) $$\nabla\cdot\vec{v} = 0\ .$$

A vector field satisfying (3.16) is sometimes called <u>solenoidal</u>.

We return now to the mathematical theory and make two remarks. First, we note that the differential operator $\nabla\cdot$ is linear in the following evident sense. Let $\vec{u},\vec{v}$ be vector fields and let a,b be real numbers. Then $a\vec{u} + b\vec{v}$ is plainly a vector field and we have

(3.17) $$\nabla\cdot(a\vec{u} + b\vec{v}) = a(\nabla\cdot\vec{u}) + b(\nabla\cdot\vec{v})\ .$$

Of course, (3.17) is what one expects from the formal properties of the inner product. The reader should notice, however, that in (3.17) a and b are, of course, constants. If a,b were non-constant scalar fields, $a\vec{u} + b\vec{v}$ would be a vector field; but (3.17) would not, in general, be true.

Next, we may consider what happens when we start with a scalar field φ, take its gradient vector field gradφ, and then apply the divergence to gradφ. The resulting scalar field is called the <u>Laplacian</u> of φ; thus the Laplacian of φ is div(gradφ) or $\nabla\cdot\nabla\varphi$. It is customary to replace $\nabla\cdot\nabla$ by the formal symbol ∇^2; and sometimes ∇^2 is itself replaced by Δ. So $\nabla^2\varphi$ and $\Delta\varphi$ are also notations for the Laplacian of the scalar field φ. From (3.1) and (3.10) we obtain immediately the formula for $\nabla^2\varphi$, (3.18) $$\nabla^2\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} \quad .$$

Example 3.3. We show that if, as in Example 3.1, we have a scalar field $\varphi(r)$, where r is the distance from the origin, then

(3.19) $$\nabla^2\varphi(r) = \varphi''(r) + \frac{2\varphi'(r)}{r} \quad .$$

First we observe that if g is a scalar field and $\vec{u}$ a vector field, then

(3.20) $$\operatorname{div}(g\vec{u}) = \vec{u}\cdot\operatorname{grad}g + g\operatorname{div}\vec{u}, \text{ or, } \nabla\cdot(g\vec{u}) = \vec{u}\cdot\nabla g + g\nabla\cdot\vec{u}.$$

Leaving the proof of (3.20) to the reader (see Section 4), we apply it to our present problem. We have already seen that

(3.21) $$\operatorname{grad}\varphi(r) = \frac{\varphi'(r)}{r}\vec{r} \ .$$

Thus

$$\nabla^2\varphi(r) = \operatorname{div}(\operatorname{grad}\varphi(r)) = \operatorname{div}\left(\frac{\varphi'(r)}{r}\vec{r}\right)$$

$$= \vec{r}\cdot\operatorname{grad}\left(\frac{\varphi'(r)}{r}\right) + \frac{\varphi'(r)}{r}\operatorname{div}\vec{r},$$

by (3.20). Reapplying (3.21), and using the evident facts that $\operatorname{div}\vec{r} = 3$ and $\vec{r}\cdot\vec{r} = r^2$, we obtain

$$\nabla^2\varphi(r) = \frac{r\varphi''(r) - \varphi'(r)}{r^3}\vec{r}\cdot\vec{r} + \frac{3\varphi'(r)}{r}$$

$$= \frac{r\varphi''(r) - \varphi'(r)}{r} + \frac{3\varphi'(r)}{r}$$

$$= \varphi''(r) + \frac{2\varphi'(r)}{r} \quad .$$

Example 3.4. In Example 2.6 we saw that the flow of heat in a solid could be represented by the vector field $\vec{q} = -K\nabla u$. Now if there are no heat sources or sinks present we can analyze the heat flow into a box, just as we did the mass flow of a fluid in Example 3.2. An entirely analogous analysis shows that the net rate at which heat flows into the box is $-\nabla\cdot\vec{q}$ per unit volume, which is equal to the rate at which the amount of heat in the box (per unit volume) is increasing, $\frac{\partial}{\partial t}(\rho c u)$, where ρ is the density, c the specific heat and u the temperature. Thus we obtain the heat equation

$$\nabla\cdot(K\nabla u) = \frac{\partial}{\partial t}(\rho c u) \ .$$

If K, ρ, c are constants this assumes the form

$$\nabla^2 u = \frac{1}{\alpha^2}\frac{\partial u}{\partial t}$$

where $\alpha^2 = \frac{K}{\rho c}$ is called the coefficient of thermal diffusivity of the material.

If steady state conditions prevail then $\frac{\partial u}{\partial t} = 0$ and the temperature satisfies Laplace's Equation $\nabla^2 u = 0$.

We turn now to the curl of $\vec{u}$. The analytical expression was given in (3.11),

$$\text{curl}\vec{u} = \nabla x\vec{u} = \left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}, \frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}, \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\right) \ .$$

The following simple hydrodynamical example illustrates the physical significance of the curl of a vector field. Consider a fluid which is rotating like a rigid body with fixed angular velocity ω about the z-axis; then the velocity vector takes the form

$$\vec{u} = \vec{\omega} \times \vec{r} = (-\omega y, \omega x, 0) \quad .$$

Then

$$\nabla x \vec{u} = (0,0,2\omega).$$

[Alternatively one could write $\nabla x(\vec{\omega} x \vec{r}) = (\nabla \cdot r)\vec{\omega} - (\vec{\omega} \cdot \nabla)\vec{r} = 3\vec{\omega} - \vec{\omega} = 2\vec{\omega}$.] The curl of the velocity vector of a fluid is known as the <u>vorticity</u> vector of the fluid; in this case, then, it is just 2ω in the direction of the axis of rotation of the fluid. In other words, the curl is twice the angular velocity vector.

The reader should have no difficulty in proving the following important properties of curl.

(3.22) $\quad \text{curl}(\text{grad}\varphi) = \vec{0}, \quad \text{or} \quad \nabla x \nabla \varphi = \vec{0}$

(3.23) $\quad \text{div}(\text{curl}\vec{u}) = 0, \quad \text{or} \quad \nabla \cdot (\nabla x \vec{u}) = 0$

(3.24) $\quad \text{curl}(a\vec{u} + b\vec{v}) = a \ \text{curl}\vec{u} + b \ \text{curl}\vec{v}$, or

$$\nabla x(a\vec{u} + b\vec{v}) = a\nabla x\vec{u} + b\nabla x\vec{v} \ ,$$

where a,b are scalar constants.

These properties are suggested by simply interpreting ∇ as a vector and applying known results of vector algebra; for example, (3.22) just reflects the fact that, for any vector

$\vec{w}$, $\vec{w}\times\vec{w} = \vec{0}$. However, the reader should be cautious about this kind of formal reasoning in dealing with formulas involving ∇. For example, (3.24) would be false if a,b were arbitrary scalar fields instead of being constant.

A vector field $\vec{u}$ such that $\mathrm{curl}\vec{u} = \vec{0}$ is called irrotational. If we adopt the language of potential theory and call a scalar field a potential, then (3.22) asserts that the gradient vector field of a potential is always irrotational. One of the most important results in the vector calculus is the converse of this, which asserts that every irrotational vector field is the gradient field of a potential. Expressing this more symbolically, we have the following theorem.

Theorem 3.2. If $\mathrm{curl}\vec{U} = 0$, then $\vec{U} = \nabla F$ for some scalar field F.

Proof. Expressing the matter in terms of components, the problem is this. We are given functions u,v,w of the variables x,y,z such that

$$\frac{\partial w}{\partial y} = \frac{\partial v}{\partial z}\ ,\ \frac{\partial u}{\partial z} = \frac{\partial w}{\partial x}\ ;\ \frac{\partial v}{\partial x} = \frac{\partial u}{\partial y}$$

and we seek a function F(x,y,z) such that

$$\frac{\partial F}{\partial x} = u,\quad \frac{\partial F}{\partial y} = v,\quad \frac{\partial F}{\partial z} = w\ .$$

We will first discuss the two-dimensional case, since this illustrates the method and will also be useful in the actual proof of the theorem. We now have to consider functions g(x,y), h(x,y) such that $\frac{\partial g}{\partial y} = \frac{\partial h}{\partial x}$ and deduce the existence of a function

$f(x,y)$ such that $\frac{\partial f}{\partial x} = g$, $\frac{\partial f}{\partial y} = h$. The key to the whole argument lies in the relation

(3.25) $$\frac{\partial^2 q}{\partial x \partial y} = \frac{\partial^2 q}{\partial y \partial x} ,$$

which holds for any function $q(x,y)$ for which the two sides of (3.25) exist and are continuous.

To solve the two-dimensional problem pick $f_1(x,y)$ so that $\frac{\partial f_1}{\partial x} = g$; then

$$\frac{\partial h}{\partial x} = \frac{\partial g}{\partial y} = \frac{\partial}{\partial y}\left(\frac{\partial f_1}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial f_1}{\partial y}\right),$$

by (3.25), and

$$\frac{\partial}{\partial x}\left(\frac{\partial f_1}{\partial y} - h\right) = 0.$$

This means that $\frac{\partial f_1}{\partial y} - h$ is a function of y only, so we may find a function $f_2(y)$ such that $\frac{df_2}{dy} = \frac{\partial f_1}{\partial y} - h$. Put $f = f_1 - f_2$. Then $\frac{\partial f_2}{\partial x} = 0$ and

$$\frac{\partial f}{\partial x} = \frac{\partial f_1}{\partial x} = g, \quad \frac{\partial f}{\partial y} = \frac{\partial f_1}{\partial y} - \frac{df_2}{dy} = h \quad .$$

Thus the two-dimensional problem is solved, and we revert to the original three-dimensional case.

Pick a function $F_1(x,y,z)$ so that $\frac{\partial F_1}{\partial z} = w$. Then $\frac{\partial u}{\partial z} = \frac{\partial w}{\partial x} = \frac{\partial}{\partial x}(\frac{\partial F_1}{\partial z}) = \frac{\partial}{\partial z}(\frac{\partial F_1}{\partial x})$, by (3.25), whence $\frac{\partial}{\partial z}(\frac{\partial F_1}{\partial x} - u) = 0$; and, similarly, $\frac{\partial v}{\partial z} = \frac{\partial w}{\partial y} = \frac{\partial}{\partial y}(\frac{\partial F_1}{\partial z}) = \frac{\partial}{\partial z}(\frac{\partial F_1}{\partial y})$, by (3.25), whence $\frac{\partial}{\partial z}(\frac{\partial F_1}{\partial y} - v) = 0$. It follows that $\frac{\partial F_1}{\partial x} - u$ is a function of x,y, say $\frac{\partial F_1}{\partial x} - u = g(x,y)$, and similarly $\frac{\partial F_1}{\partial y} - v = h(x,y)$. Then $\frac{\partial g}{\partial y} = \frac{\partial}{\partial y}\ (\frac{\partial F_1}{\partial x} - u) = \frac{\partial^2 F_1}{\partial y \partial x} - \frac{\partial u}{\partial y} = \frac{\partial^2 F_1}{\partial x \partial y} - \frac{\partial v}{\partial x} = \frac{\partial h}{\partial x}$. We are thus back in the two-dimensional case and may infer the existence of a function $F_2(x,y)$ such that $\frac{\partial F_2}{\partial x} = g$, $\frac{\partial F_2}{\partial y} = h$.

Put $F = F_1 - F_2$. Then $\frac{\partial F_2}{\partial z} = 0$, and

$$\frac{\partial F}{\partial x} = \frac{\partial F_1}{\partial x} - \frac{\partial F_2}{\partial x} = \frac{\partial F_1}{\partial x} - g = u\ ,$$

$$\frac{\partial F}{\partial y} = \frac{\partial F_1}{\partial y} - \frac{\partial F_2}{\partial y} = \frac{\partial F_1}{\partial y} - h = v\ ,$$

$$\frac{\partial F}{\partial z} = \frac{\partial F_1}{\partial z} - \frac{\partial F_2}{\partial z} = \frac{\partial F_1}{\partial z} = w.$$

This proves the theorem. Notice that in fact we may only alter F by an arbitrary constant, since the difference of any two choices of F would have to be a function all of whose partial derivatives vanish.

In solving specific problems it is often convenient to use a systematic procedure such as is indicated in the following example.

Example 3.5. Show that the vector field $\vec{u} = (1 + 2xy + z, x^2 + 2y + z, x + y + z)$ is irrotational, and find a scalar field F such that $\mathrm{grad} F = \vec{u}$.

Here $\mathrm{curl}\vec{u} = (1-1, 1-1, 2x-2x) = \vec{0}$, so $\vec{u}$ is irrotational. Thus by Theorem 3.2 there is a function $F(x,y,z)$ such that

$$\frac{\partial F}{\partial x} = 1 + 2xy + z, \tag{3.26}$$

$$\frac{\partial F}{\partial y} = x^2 + 2y + z, \tag{3.27}$$

$$\frac{\partial F}{\partial z} = x + y + z\ . \tag{3.28}$$

Equation (3.26) will be satisfied if F has the form

$$F = x + x^2 y + xz + G(y,z)\ . \tag{3.29}$$

Here the function $x + x^2 y + xz$ is obtained by integrating the right hand side of (3.26) with respect to x, regarding y and z as constants. $G(y,z)$ can be any function of y and z as far as satisfying (3.26) is concerned, but we must now choose it to satisfy (3.27) and (3.28).

Substituting (3.29) in (3.27) gives

$$x^2 + \frac{\partial G}{\partial y} = x^2 + 2y + z\ , \tag{3.30}$$

or

$$\frac{\partial G}{\partial y} = 2y + z. \tag{3.31}$$

Here we have a partial check on our work thus far. Since G does not involve x neither can $\frac{\partial G}{\partial y}$; hence in passing from (3.30) to (3.31) the terms involving x must cancel. If they do not we have evidence of something wrong. Either $\text{curl}\vec{u} \neq \vec{0}$ or a mistake has been made in the algebra leading up to (3.31).

Integrating (3.31) with respect to y, holding z constant, gives

$$G = y^2 + yz + H(z) \ ,$$

and so (3.29) becomes

$$F = x + x^2y + xz + y^2 + yz + H. \tag{3.32}$$

We now substitute this into the third equation (3.28) to get, after simplifying,

$$\frac{dH}{dz} = z. \tag{3.33}$$

Here we have another check on the algebra, since the right hand side of (3.33) must reduce to a function of z only.

A final integration gives

$$H = \tfrac{1}{2}z^2 + c \ ,$$

and so we get as our solution,

$$F = x^2 + x^2y + xz + y^2 + yz + \tfrac{1}{2}z^2 + c,$$

where c is any constant.

The solution of such a problem should always be checked by computing gradF and seeing that it actually does equal the given $\vec{u}$.

Problems

3.1 Find the tangent planes and normals to the surfaces given below at the points named.

(a) $z = x^2 + y^2$; $(1,2,5)$.

(b) $x^2 + y^2 + z^2 - 4x + 6y - 2z + 3 = 0$; $(3,-2,4)$.

(c) $x = \lambda \cos\mu$, $y = \lambda \sin\mu$, $z = \lambda\mu$; $\lambda = 1$, $\mu = \frac{\pi}{2}$.

3.2 Find the divergence of the following vector fields.

(a) (x,y,z).

(b) (y,z,x).

(c) $(x^2 + y^2, xy, y^2 + z^2)$.

(d) $\vec{r}$.

(e) $\frac{\vec{r}}{\|\vec{r}\|^2}$. [Hint. Apply (3.20) to $r^{-2}\vec{r}$.]

3.3 Find the curl of the vector fields in Problem 3.2. [Hint, (e). Show first that $r^{-2}\vec{r} = \nabla \log r$.]

3.4 Prove (3.21), (3.22), (3.23).

3.5 Show that $\operatorname{div}(\vec{u}\times\vec{v}) = (\operatorname{curl}\vec{u})\cdot\vec{v} - (\operatorname{curl}\vec{v})\cdot\vec{u}$.

3.6 Show that, if φ is a scalar field and u a vector field, then $\operatorname{div}(\varphi\vec{u}) = \varphi \operatorname{div}\vec{u} + \operatorname{grad}\varphi\cdot\vec{u}$ and obtain a similar formula for $\operatorname{curl}(\varphi\vec{u})$. (See (3.20)).

3.7 Show that each of the following vector fields is irrotational, and find potentials of which they are the gradients.

(a) $\frac{x}{x^2+y^2+z^2}$, $\frac{y}{x^2+y^2+z^2}$, $\frac{z}{x^2+y^2+z^2}$. (cf. Problem 3.3(e))

(b) $(2x + y + 2, x + z + 3, y)$

(c) $(\cos x \cos y - \sin x \sin z, \cos y \cos z - \sin y \sin x, \cos z \cos x - \sin z \sin y)$.

3.8 Show that $r^n\vec{r}$ is irrotational. Is it solenoidal?

3.9 For what values of n does r^n satisfy Laplace's Equation (for $r \neq 0$)(use (3.19)).

3.10 Let $\vec{\omega} = w\vec{k}$, w being constant, and let $\vec{v} = \frac{1}{r^2}\,\vec{\omega}\times\vec{r}$ Considering only points in the xy-plane, i.e., $\vec{r} = x\vec{i} + y\vec{j}$, show that $\vec{v}$ is solenoidal and irrotational. Find a potential function for $\vec{v}$. Answer. Potential = $-w \arctan(x/y)$, or $w \arctan(y/x)$.

3.11 A fluid is in radial outward flow from a spherical source, with velocity of constant magnitude v_0. Assuming that the density ρ is a function of r only, find this function. Answer. $\rho = c/r^2$.

3.12 Consider an inviscous fluid whose velocity at point (x,y,z) at time t is given by $\vec{v} = v_1(x,y,z,t)\vec{i} + v_2(x,y,z,t)\vec{j} + v_3(x,y,z,t)\vec{k}$ and whose density is $\rho(x,y,z,t)$ and pressure $p(x,y,z,t)$. Suppose that a body force $\vec{F}(x,y,z,t)$ per unit of mass is also present.

(a) Applying Newton's law of motion show that

$$\rho\,\frac{d\vec{v}}{dt} = \vec{F}\rho - \nabla p\ ,$$

(b) Using the fact that

$$\frac{d\vec{v}}{dt} = \left(\frac{\partial v_1}{\partial x}\,\frac{dx}{dt} + \frac{\partial v_1}{\partial y}\,\frac{dy}{dt} + \frac{\partial v_1}{\partial z}\,\frac{dz}{dt} + \frac{\partial v_1}{\partial t}\right)\vec{i} + \text{etc.}$$

$$= \left(v_1\,\frac{\partial v_1}{\partial x} + v_2\,\frac{\partial v_1}{\partial y} + v_3\,\frac{\partial v_1}{\partial z} + \frac{\partial v_1}{\partial t}\right)\vec{i} + \text{etc.}$$

derive Euler's Equations

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v}\cdot\nabla)\vec{v} = \vec{F} - \frac{1}{\rho}\nabla p \ .$$

(c) This vector equation is equivalent to three scalar equations. The continuity equation (3.14) $\nabla\cdot(\rho\vec{v}) = -\frac{\partial\rho}{\partial t}$ provides a fourth equation. There are five unknown functions v_1, v_2, v_3, ρ, p. One usually wants as many equations as there are unknowns. From what physical considerations might the fifth relation be derivable?

3.13 A first order differential equation written in the form $M(x,y)dx + N(x,y)dy = 0$ is said to be <u>exact</u> if $\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x} = 0$. Show that the equation is exact if and only if there is a function $F(x,y)$ such that $Mdx+Ndy = dF$. A solution of the differential equation is then $F(x,y) = C$, where C is any constant.

3.14 A function $Q(x,y)$ is called an <u>integrating factor</u> (cf. Chapter 1, Section 6) of the equation $M(x,y)dx + N(x,y)dy = 0$ if $QMdx + QNdy = 0$ is exact. Show that $Mdx + Ndy = 0$ has an integrating factor $Q(x)$ depending only on x if and only if the quantity $R = (\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x})/N$ is independent of y. If this condition is satisfied show that $Q = e^{\int R(x)dx}$.

3.15 Apply the results of Problem 3.14 to solve the equation

$$(3x + 2y^2)dx + 2xydy = 0.$$

3.16 Suppose the temperature in a body is given by

$$u(x,y,z) = xy^2 + yz^2 + zx^2$$

in degrees centigrade, where x,y,z are measured in centimeters.

Consider a small square of area 0.0001 cm^2 lying in a horizontal plane at the point P(1,2,3). (The z-axis is ↑).

(a) Compute the temperature gradient at the point P.

(b) If the coefficient of thermal conductivity is 0.2 cal/cm-sec°C what is the rate at which heat flows across the square? Does it flow from the top face of the square to the bottom face, or the other way?

Ans. (10,13,13); 2.6 × 10-5 cal/sec.

3.17 Given the point P(1,2,3), the scalar field $f = yz + zx + xy$, and the vector fields $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$ and $\vec{v} = yz\vec{i} + zx\vec{j} + xy\vec{k}$, compute the following:

(a) gradf at P;

(b) directional derivative of f in the direction of $\vec{v}$ at P;

(c) curl$\vec{v}$;

(d) $\nabla \cdot (\vec{r} \times \vec{v})$;

(e) Laplacian of f;

(f) the value of the constant c for which the surface $f = c$ passes through P;

(g) the angle between $\vec{r}$ and the normal to the surface $f = c$ at P;

(h) the equation of the plane tangent to the surface $f = c$ at P.

3.18 The surfaces $x^2 + y^2 + z^2 = 14$ and $xyz = 6$ intersect at the point (1,2,3). Find the angle at which they intersect (i.e., the angle between their normals) at this point.

3.19 Let the vector field $\vec{F} = 2xy\vec{i} + (x^2 + z^3)\vec{j} + 3yz^2\vec{k}$

(a) Compute the divergence of $\vec{F}$.

(b) Find the curl of $\vec{F}$.

(c) Which of the above parts implies that there exists a scalar potential field φ such that $\vec{F} = \nabla\varphi$?

(d) Find the φ of part (c).

3.20 Suppose the temperature in a body is given by

$$u(x,y,z) = xy + yz^2 \ .$$

Suppose a particle is moving through the body along the path

$$\vec{r} = ti + t^2j + t^3k \ .$$

Thus when $t = 2$ the particle is at the point $P = (2,4,8)$. Find the directional derivative of the temperature at the point P in the direction of the velocity of the particle as it passes through P. Answer: $1036/\sqrt{161}$.

3.21 Write the equation of the line tangent to the curve of intersection of the surfaces $xy + yz + zx = 11$ and $x^3y^2z = 12$ at the point $(1,2,3)$. Answer: $\vec{r} = (1-5t)\vec{i} + (2 + 22t)\vec{j} + (3 - 21t)\vec{k}$.

3.22 (a) Show that the curve $\vec{r} = (t^3, 2t, 3t^2)$ intersects the surface $x^2 - 2xy + 2yz = 9$ at the point $P(1,2,3)$.

(b) Find the angle at P between the tangent to the curve and the normal to the surface.

(c) Find the projection of the <u>unit</u> tangent to the curve at P into the plane tangent to the surface at P.

Answer: (b) $\theta = \arccos 13/21$. (c) $4/63(10,-2,7)$.

4. Algebra of the Operator ∇.

In the earlier sections of this chapter we found it useful to introduce the vector differential operator

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right),$$

and we have seen that it can often, but not always, be manipulated as if it were an ordinary vector. It is worthwhile to examine this situation more carefully, to see, if possible, just which procedures can be carried out and which cannot.

Let us first take a look at the more familiar differential operator $D = \frac{d}{dx}$. Its basic algebraic properties are embodied in the formulas

(4.1) $D(f+g) = Df+Dg$, $D(cf) = c(Df)$, $D(fg) = f(Dg)+(Df)g$,

for any differentiable functions $f(x)$, $g(x)$ and any constant c.

Corresponding to the first two parts of (4.1) we have the linearity properties of the gradient, the divergence, and the curl, which have already been mentioned. The second part of (4.1) also leads at once to the identities

(4.2) $$\nabla\cdot(\varphi\vec{c}) = (\nabla\varphi)\cdot\vec{c}, \quad \nabla\times(\varphi\vec{c}) = (\nabla\varphi)\times\vec{c},$$

if $\vec{c}$ is a constant vector.

Now consider $\nabla\cdot(\vec{c}\times\vec{v})$; here it is not so obvious what to do. The result must be a scalar and must involve $\vec{c}$ and ∇ operating on $\vec{v}$, but how do we put these features together? One way of proceeding is to forget about the special character of ∇ and treat it just like an ordinary vector, manipulate the expression algebraically until it has the form we want, and then check the result to see if we really have an identity. Thus with our present problem we can proceed as follows:

$$\nabla\cdot(\vec{c}\times\vec{v}) = [\nabla,\vec{c},\vec{v}] = -[\vec{c},\nabla,\vec{v}] = -\vec{c}\cdot(\nabla\times\vec{v}).$$

Now we check:

$$\nabla\cdot(\vec{c}\times\vec{v}) = \frac{\partial}{\partial x}(c_2v_3-c_3v_2) + \frac{\partial}{\partial y}(c_3v_1-c_1v_3) + \frac{\partial}{\partial z}(c_1v_2-c_2v_1)$$

$$-\vec{c}\cdot(\nabla\times\vec{v}) = -c_1\left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right) - c_2\left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right) - c_3\left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right).$$

Comparison shows that indeed

(4.3) $$\nabla\cdot(\vec{c}\times\vec{v}) = -\vec{c}\cdot(\nabla\times\vec{v}) .$$

The combination $\nabla\times(\vec{c}\times\vec{v})$ is still harder to handle. There is no way to manipulate such a triple vector product without using Theorem 1.9 to break it up. If we do this, purely formally, we get

$$\nabla\times(\vec{c}\times\vec{v}) = (\nabla\cdot\vec{v})\vec{c} - (\nabla\cdot\vec{c})\vec{v} .$$

Here the last term is obviously at fault, since we want ∇ to operate on $\vec{v}$, not on $\vec{c}$. But this term can be written, again purely formally, as $(\vec{c}\cdot\nabla)\vec{v}$, and this can be interpreted as an operator, $\vec{c}\cdot\nabla$, operating on $\vec{v}$. Thus we get finally

(4.4) $$\nabla\times(\vec{c}\times\vec{v}) = (\nabla\cdot\vec{v})\vec{c} - (\vec{c}\cdot\nabla)\vec{v}.$$

It is worth saying a word about the operator $\vec{c}\cdot\nabla$. This is a scalar differential operator, and so can be applied to a scalar function to get another scalar function or, as in (4.4), to a vector function to get another vector function. If $\vec{c}$ is a unit vector, $\vec{c}\cdot\nabla$ is just the derivative in the direction $\vec{c}$, as one sees in Theorem 2.1. An important physical appearance of such an operator is in the Euler Equations of fluid flow (see Problem 3.12b).

So far we have avoided the analogs of the third part of (4.1), the product rule. Some vector applications are fairly straightforward, such as

$$\begin{aligned}\nabla(\varphi\psi) &= \varphi(\nabla\psi) + (\nabla\varphi)\psi,\\ \nabla\cdot(\varphi\vec{v}) &= \varphi(\nabla\cdot\vec{v}) + (\nabla\varphi)\cdot\vec{v}\ ,\\ \nabla\mathrm{x}(\varphi\vec{v}) &= \varphi(\nabla\mathrm{x}\vec{v}) + (\nabla\varphi)\mathrm{x}\vec{v}\ ,\end{aligned} \tag{4.5}$$

(see Problem 3.6). But other cases require more elaborate techniques. One very useful device is to regard D, when operating on a product fg, as consisting of two parts, $D = D_f + D_g$, where D_f acts only on f and D_g only on g; more precisely, $D_f g = 0$ and $D_g f = 0$. Then

$$D(fg) = (D_f+D_g)(fg) = D_f(fg)+D_g(fg) = (Df)g+f(Dg)$$

as desired. If we apply the same trick to ∇ we get, for example,

$$\begin{aligned}\nabla\cdot(\vec{u}\mathrm{x}\vec{v}) &= \nabla_u\cdot(\vec{u}\mathrm{x}\vec{v}) + \nabla_v\cdot(\vec{u}\mathrm{x}\vec{v})\\ &= [\nabla_u,\vec{u},\vec{v}] + [\nabla_v,\vec{u},\vec{v}]\\ &= [\nabla_u,\vec{u},\vec{v}] - [\vec{u},\nabla_v,\vec{v}]\\ &= (\nabla\mathrm{x}\vec{u})\cdot\vec{v} - \vec{u}\cdot(\nabla\mathrm{x}\vec{v})\end{aligned} \tag{4.6}$$

This can, and should, be checked for correctness. (See Problem 3.5.)

We can consider many other complications, but the ideas introduced above will handle most of the situations that are likely to arise, including those involving more than one occurrence of the operator ∇. If we wish we can push the algebra still farther and bring in matrices. As in Chapter 5 let our vectors, including ∇, be column vectors. Then we can write $\nabla\cdot\vec{v}$ as $\nabla^t\vec{v}$. On the other hand $\nabla\vec{v}^t$ is the 3x3 matrix

$$\begin{pmatrix} \frac{\partial v_1}{\partial x} & \frac{\partial v_2}{\partial x} & \frac{\partial v_3}{\partial x} \\ \frac{\partial v_1}{\partial y} & \frac{\partial v_2}{\partial y} & \frac{\partial v_3}{\partial y} \\ \frac{\partial v_1}{\partial z} & \frac{\partial v_2}{\partial z} & \frac{\partial v_3}{\partial z} \end{pmatrix}.$$

The transpose $(\nabla\vec{v}^t)^t$ of this matrix is called the Jacobian matrix of $\vec{v}$; it has many uses in the more extensive study of vector fields. (Note that we cannot write $(\nabla\vec{v}^t)^t = \vec{v}\nabla^t$. This last expression is a matrix whose elements are operators.) The matrix operator

$$\nabla\nabla^t = \begin{pmatrix} \frac{\partial^2}{\partial x^2} & \frac{\partial^2}{\partial x\partial y} & \frac{\partial^2}{\partial x\partial z} \\ \frac{\partial^2}{\partial y\partial x} & \frac{\partial^2}{\partial y^2} & \frac{\partial^2}{\partial y\partial z} \\ \frac{\partial^2}{\partial z\partial x} & \frac{\partial^2}{\partial z\partial y} & \frac{\partial^2}{\partial z^2} \end{pmatrix}$$

(called the Hessian operator) arises in the theory of maxima and minima of scalar functions of three variables.

Problems

4.1 (a) By expanding $\vec{c} \times (\nabla \times \vec{v})$ derive the identity $\nabla(\vec{c} \cdot \vec{v}) = \vec{c} \times (\nabla \times \vec{v}) + (\vec{c} . \nabla)\vec{v}$.

(b) Using the above result, and expressing ∇ as $\nabla_u + \nabla_v$, derive a formula for $\nabla(\vec{u} \cdot \vec{v})$.

4.2 There are five triple products involving the operator ∇ twice, namely $\nabla \cdot (\nabla \varphi)$, $\nabla \times (\nabla \varphi)$, $\nabla \cdot (\nabla \times \vec{v})$, $\nabla \times (\nabla \times \vec{v})$, $\nabla(\nabla \cdot v)$.

Show the following:

(a) $\nabla \cdot (\nabla \varphi)$ has been discussed and has a name. What is it?

(b) $\nabla \times (\nabla \varphi)$ is identically zero.

(c) $\nabla \cdot (\nabla \times \vec{v})$ is identically zero.

(d) $\nabla \times (\nabla \times \vec{v}) = \nabla(\nabla \cdot \vec{v}) - \nabla^2 \vec{v}$.

(e) $\nabla(\nabla \cdot \vec{v})$ is the Hessian of $\vec{v}$.

4.3 One might argue that $\nabla \cdot (\nabla \times \vec{v}) = 0$ because this is a triple scalar product with two equal factors. By the same argument $\vec{v} \cdot (\nabla \times \vec{v}) = 0$. Is this true?

4.4 Derive and check a formula for $\nabla \times (\vec{u} \times \vec{v})$.

4.5 Show that

$\text{curl}^3 \vec{v} = \text{curl}(\text{curl}(\text{curl}\vec{v})) = -\nabla^2(\text{curl}\vec{v})$

Generalize to

$\text{curl}^{2n+1} \vec{v} = (-\nabla^2)^n (\text{curl}\vec{v})$.

4.6 (a) If $\vec{w}$ is a constant vector and ϕ a function of r only, ($r = \|\vec{r}\|$), prove the identity

$$\nabla x(\phi\vec{w}) = \frac{1}{r}\frac{d\phi}{dr}\vec{r} \times \vec{w}.$$

(b) Use this identity to solve the following problem: Given any constant vector $\vec{w}$, find a vector field $\vec{v}$ such that $\nabla x\vec{v} = \vec{w} x\vec{r}$.

4.7 If $\vec{r}$ is the radius vector (x,y,z) show that $(\vec{v}\cdot\nabla)\,\vec{r} = \vec{v}$ and $\nabla\cdot\vec{r} = 3$. Then show that for any constant vector $\vec{\omega}$, $\text{curl}(\vec{\omega} x\vec{r}) = 2\vec{\omega}$.

Interpret this in terms of rotation about an axis.

5. Change of Coordinates.

So far we have been assuming throughout this chapter that E_3 is furnished with a fixed right-handed Cartesian coordinate system. In practice, however, we often wish to be free to choose our coordinate system to suit the particular problem under discussion; and there are many problems for which the most appropriate coordinates are not Cartesian at all, but, say, cylindrical polar coordinates or spherical polar coordinates. However, our definitions in the previous sections and, in particular, our definition of the differential operator ∇, have all been made in terms of a fixed coordinate system . From those definitions it is not at all clear whether the basic concepts (grad, div, curl, Laplacian) are, in a sense to be made precise, independent of the coordinate system given We referred briefly to this matter in the previous section, in our discussion of the geometrical and physical significance of div and curl. We take it up again now in greater detail; but we do not aim at a complete treatment, being content to explain the precise nature of the <u>invariance problem</u>, as it is often called.

At first we will confine our attention to Cartesian coordinate systems. As in Chapter 2, Section 15 and Chapter 5, Section 7 we must distinguish between a physical vector $\vec{v}$ and the number triple that represents it in a given basis. If $\{\vec{i},\vec{j},\vec{k}\}$ and $\{\vec{i}',\vec{j}',\vec{k}'\}$ are two right-handed orthonormal bases, and

$$\vec{v} = x_1\vec{i}+x_2\vec{j}+x_3\vec{k} = x_1'\vec{i}'+x_2'\vec{j}'+x_3'\vec{k}' ,$$

then the two <u>different</u> triples (x_1,x_2,x_3) and (x_1',x_2',x_3') represent the <u>same</u> physical vector $\vec{v}$ in the two different bases. (In more technical terms, we have two different isomorphisms between the space V of physical vectors and the space V_3 of number triples.)

In Chapter 2 (see, in particular, equation (15.9)) we showed that in this situation there is a non-singular matrix C, determined by the relation between the two bases, such that

$$(5.1) \qquad \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = C \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} .$$

This relation can also be written in the form

$$(5.2) \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = B \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} ,$$

where B is the inverse of C, $B = C^{-1}$. Now if we apply (5.2) to the vector $\vec{i}'$, whose x' coordinates are $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, we get

$$B \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix} .$$

Treating $\vec{j}'$ and $\vec{k}'$ similarly, we see that the column vectors of B are just the representations, in the $\vec{i},\vec{j},\vec{k}$ system of the $\vec{i}',\vec{j}',\vec{k}'$ vectors. Now these three vectors form an orthonormal system, and so the column vectors of B form an orthonormal systen in the space of column vectors. It then follows from Theorem 8.2 of Chapter 5 that B is an orthogonal matrix, and as a simple consequence of Corollaries 8.1 and 8.2 that C is orthogonal. In fact, $B = C^t$, so that $b_{ij} = c_{ji}$.

Now from (5.1), which we write in the more convenient form $x' = Cx$, and a similar equation $y' = Cy$ for some other vector, we get

$$\Sigma x_i' y_i' = x'^t y' = (Cx)^t (Cy) = x^t C^t C y = x^t y = \Sigma x_i y_i ,$$

since $C^t C = I$ because C is orthogonal. That is, we have proved that <u>the algebraic definition (1.1) of the inner product is invariant under change of orthonormal bases</u>. It is true that it is not imperative that we have such an algebraic proof, since we have shown the invariance of the inner product by a geometric argument. However, as we shall see, such a purely algebraic argument is useful in cases where a geometric one may not be available.

A similar but more complicated algebraic proof can be given for the invariance of the cross product. We shall not carry this out in detail, but there is one additional feature of the matrix C that must be noted. If the algebraic definition (1.4) of the cross product is to be invariant then, in particular, the algebraic expression in Theorem 1.7 must hold for $(\vec{i} \times \vec{j}) \cdot \vec{k} = 1$ when these vectors are expressed in terms of the $\{\vec{i}', \vec{j}', \vec{k}'\}$ basis. By an argument like one used above we see that the primed coordinates of the $\vec{i}, \vec{j}, \vec{k}$ vectors are the column vectors of C. It follows from Theorem 1.7 that $(\vec{i} \times \vec{j}) \cdot \vec{k}$ $= \det C^t = \det C$. Hence <u>we must restrict our coordinate transformations to those for which $\det C = 1$</u>. It can be shown (see Problem 5.6) that C is then actually a <u>rotation</u> of coordinates. The two bases, $\{\vec{i}, \vec{j}, \vec{k}\}$ and $\{\vec{i}', \vec{j}', \vec{k}'\}$ have the same orientation, i.e., if one of them is right-handed so is the other.

We have now completed the discussion of change of coordinates. We have seen that the allowed changes of (Cartesian) coordinates are all effected by orthogonal transformations of determinant +1, and we have the analytical formula (5.1) for a typical change of coordinates. We recall that the origin of coordinates remains fixed under the changes of coordinates we are considering and that all coordinate systems so far considered are right-handed Cartesian systems. **We now take up the problem of invariance.**

Suppose that φ is a real valued function of points of E_3; for example, $\varphi(P)$ might be the distance OP from the origin to P, or $\varphi(P)$ might be the temperature at the point P. If we choose a coordinate system in E_3, then φ is represented by a function $f(x,y,z)$ of the three real variables (x,y,z); that is, we define f by

$$f(x,y,z) = \varphi(P) , \tag{5.3}$$

where (x,y,z) are the coordinates of P. We may now form the vector field gradf in V_3. This associates with every point (x_0,y_0,z_0) of E_3 an element of V_3, given by

$$(\text{grad}f)_0 = \left(\frac{\partial f}{\partial x} , \frac{\partial f}{\partial y} , \frac{\partial f}{\partial z}\right)_0 , \tag{5.4}$$

the zero subscript indicating that the quantities are evaluated at the point (x_0,y_0,z_0). Now through the coordinatization, the vector on the right of (5.4) corresponds to some directed line segment OQ_0 in E_3. Thus if (x_0,y_0,z_0) are the coordinates of P_0 we may <u>define</u> gradφ by the rule

(5.5) $$(\text{grad}\varphi)(P_0) = 0Q_0 .$$

All that we have done so far is to reinterpret grad in terms of E_3 itself rather than its coordinatization. However the definition of gradφ has been through the medium of a certain fixed coordinate system and the question of invariance is just this: if we choose a different coordinate system do we change the definition of gradφ?

Now in Section 3 we showed that a change of coordinate system does not in fact change the definition of gradφ; this we did by giving an invariant characterization of gradφ which made no reference to coordinates. Namely, we showed that the direction of gradφ is the direction of maximum rate of change of φ, and the magnitude $\|\text{grad}\varphi\|$ is the maximum rate of change of φ. This gives us a determination of the vector field gradφ quite independent of coordinates. Thus one procedure for proving that gradφ, div$\vec{u}$, curl$\vec{u}$, and so on are invariant (i.e., independent of choice of coordinate system) is to characterize them in a way which is plainly invariant. Naturally, we can readily convince ourselves of the invariance of these operations by considering the physical situations in which they arise; since, if they occur naturally in the dynamics of bodies in 3-space, they must be independent of the choice of coordinate systems introduced for the purpose of giving a convenient analytical formula for them.

However, a systematic mathematical procedure is also available for invariance proofs, since we know how to transform from one Cartesian coordinate system to another. This procedure has the added merit of being valid even when we have no reliable physical intuition to guide us.

To make most efficient use of matrix methods we shall use x_1, x_2, x_3 as coordinates instead of x, y, z. Then we have

$$\vec{r} = x_1\vec{i}+x_2\vec{j}+x_3\vec{k} = x_1'\vec{i}'+x_2'\vec{j}'+x_3'\vec{k}'$$

and so (5.1) applies, thus

$$(5.1) \qquad \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = C \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} .$$

Since C is orthogonal, $C^t = C^{-1}$, and so this equation can be written as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = C^{-1} \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = C^t \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} .$$

Written in terms of elements this is

$$x_j = \sum_{i=1}^{3} c_{ij}x'_i, \quad j = 1,2,3.$$

From this we get

(5.6) $$\frac{\partial x_j}{\partial x_i'} = c_{ij}, \quad i,j = 1,2,3 ,$$

a relation that we shall use later.

We suppose that the scalar field φ is represented in the one coordinate system by the function f and in the other by the function f', so that

(5.7) $$f(x_1,x_2,x_3) = f'(x_1',x_2',x_3')$$

where the coordinates are related by (5.6). Let us write gradf as the column vector

$$\begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} f .$$

Then we wish to show that if $\begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix}$ f is evaluated at (x_1,x_2,x_3)

and $\begin{pmatrix} \frac{\partial}{\partial x_1'} \\ \frac{\partial}{\partial x_2'} \\ \frac{\partial}{\partial x_3'} \end{pmatrix}$ f' is evaluated at (x_1',x_2',x_3') then the (vector) values are themselves related through the relationship (5.1). That is, we wish to show that

$$(5.8)\qquad \begin{pmatrix} \dfrac{\partial}{\partial x_1{}'} \\[2ex] \dfrac{\partial}{\partial x_2{}'} \\[2ex] \dfrac{\partial}{\partial x_3{}'} \end{pmatrix} f' = C \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\[2ex] \dfrac{\partial}{\partial x_2} \\[2ex] \dfrac{\partial}{\partial x_3} \end{pmatrix} f,$$

where the left-hand side is evaluated at $(x_1{}',x_2{}',x_3{}')$ and the right-hand side at (x_1,x_2,x_3). Now the chain-rule (2.6) tells us that

$$\frac{\partial f'}{\partial x_i{}'} = \sum_{j=1}^{3} \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial x_i{}'} = \sum_{j=1}^{3} \frac{\partial f}{\partial x_j} c_{ij},$$

the last equality following from (5.6). Writing this as

$$\frac{\partial f'}{\partial x_i{}'} = \sum_{j=1}^{3} c_{ij} \frac{\partial f}{\partial x_j} ,$$

we see that it can be put in the matrix form

$$(5.9)\qquad \begin{pmatrix} \dfrac{\partial}{\partial x_1{}'} \\[2ex] \dfrac{\partial}{\partial x_2{}'} \\[2ex] \dfrac{\partial}{\partial x_3{}'} \end{pmatrix} f' = C \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\[2ex] \dfrac{\partial}{\partial x_2} \\[2ex] \dfrac{\partial}{\partial x_3} \end{pmatrix} f \quad .$$

The argument given above is perfectly rigorous. In it we have come very close to treating ∇ as a vector. By adopting a suitable convention we may indeed write (5.9) in the symbolic form

$$(5.10)\qquad \begin{pmatrix} \frac{\partial}{\partial x_1{}'} \\ \frac{\partial}{\partial x_2{}'} \\ \frac{\partial}{\partial x_3{}'} \end{pmatrix} = C \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} .$$

where it is now understood that the two differential operators operate on functions $\varphi(P)$ of points in E_3. Formula (5.10) suggests immediately a method of proof of the invariance of div and curl. For if a vector field $\vec{U}$ is given over E_3 which corresponds in the two coordinate systems to vector fields $\vec{u}$ and $\vec{u}'$, then

$$(5.11)\qquad \vec{u}' = C\vec{u} ,$$

where $\vec{u}$ is evaluated at (x_1, x_2, x_3) and $\vec{u}'$ at $(x_1{}', x_2{}', x_3{}')$. Writing ∇, ∇' for the differential operators in (5.10), we may rewrite that formula as

$$(5.12)\qquad \nabla' = C\nabla .$$

Formulas (5.11) and (5.12) immediately yield the invariance. For $\nabla' \cdot \vec{u}' = C\nabla \cdot C\vec{u} = \nabla \cdot \vec{u}$, since the inner product is invariant under orthogonal transformations. Also, if C is a rotation matrix, i.e., if $\det C = +1$, then $\nabla' x \vec{u}' = C\nabla x C\vec{u} = \nabla x \vec{u}$, since the cross product is invariant under rotation. Thus the invariance of the divergence and curl simply reflect the geometrical properties of the inner and cross products, and the treatment of ∇ as a vector, at least in the arguments given, has been seen to be justified.

Of course, the invariance of the Laplacian $\nabla^2\varphi$ follows immediately from that of grad and divergence.

We now take up, rather briefly, the question of the formulas for the differential operators discussed in non-Cartesian coordinate systems. This is very important since, in many problems and many physical situations (e.g., spherical symmetry), other coordinate systems are more appropriate. We will content ourselves with dealing with the Laplacian in cylindrical and spherical polar coordinates and just stating the facts. First, however, we should make a brief remark to justify our concentration on the Laplacian.

The equation $\vec{u} = \text{grad}\varphi$ (or* $\vec{u} = -\text{grad}\varphi$) occurs frequently in various physical theories. Thus φ may be the gravitational potential, the electrostatic potential, the magnetic potential, temperature, pressure, density and so on; and $\vec{u}$ is then the corresponding gradient vector field. The equation, $\nabla\cdot\vec{u} = 4\pi\sigma$, relates the velocity vector of an incompressible fluid to the source density. If the fluid flow is irrotational then, by Theorem 3.2, $\vec{u} = -\text{grad}\varphi$ for some velocity potential φ, and so

(5.13) $$\nabla^2\varphi = -4\pi\sigma \quad .$$

*We saw from examples in Section 2 that we might also be interested in the negative of the gradient of a scalar field.

However we may also interpret $\vec{u}$ as the electric intensity in free space and we again get equation (5.13) with φ now standing for electrostatic potential. Again we may suppose $\vec{u}$ to be the gravitational force vector and that σ represents the density of matter. We would then write $\nabla \cdot \vec{u} = -4\pi\sigma$ since a material particle is regarded as a sink; and if $\vec{u}$ is irrotational and $\vec{u} = \text{grad}\varphi$, we again get equation (5.13). This important equation is usually called Poisson's equation; a very important special case is that in which there is no source distribution, so that (5.13) takes the form

$$\nabla^2\varphi = 0 , \tag{5.14}$$

known as Laplace's equation.

These remarks should suggest the great importance of the Laplacian operator ∇^2. The partial differential equation (5.14) has been studied extensively by physicists, engineers, and mathematicians under the assumption of certain specific boundary conditions. It is of interest in the 2-dimensional case, as well as in three dimensions.

We have already seen that, under the assumption of spherical symmetry we have $\varphi = \varphi(r)$ and

$$\nabla^2\varphi = \varphi'' + \frac{2\varphi'}{r} . \tag{5.15}$$

It is worth noting that, in this case, the solution of Laplace's equation involves only the differential equation

$$\varphi'' + \frac{2\varphi'}{r} = 0 ,$$

from which we easily deduce that $\varphi = A + \frac{B}{r}$. This then gives the form of the spherically symmetric potential in the absence of source.

The general expression for the Laplacian in spherical polar coordinates (r,θ,ψ) is

$$(5.16) \quad \nabla^2\varphi = \frac{\partial^2\varphi}{\partial r^2} + \frac{2}{r}\frac{\partial\varphi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\varphi}{\partial\theta^2} + \frac{\cot\theta}{r^2}\frac{\partial\varphi}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\varphi}{\partial\psi^2} .$$

This may be deduced directly (but tediously) from the chain-rule, or it may be inferred from the formula for ∇ with respect to the local coordinate system given by (r,θ,ψ), namely

$$\nabla = \left(\frac{\partial}{\partial r} , \frac{1}{r}\frac{\partial}{\partial\theta} , \frac{1}{r\sin\theta}\frac{\partial}{\partial\psi}\right) .$$

Similarly in cylindrical polar coordinates (r,θ,z) we have

$$(5.17) \qquad \nabla^2\varphi = \frac{\partial^2\varphi}{\partial r^2} + \frac{1}{r}\frac{\partial\varphi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\varphi}{\partial\theta^2} + \frac{\partial^2\varphi}{\partial z^2} .$$

In this case the formula may be inferred immediately from that for the 2-dimensional Laplacian in polar coordinates. [Note that in (5.17), and in Problem 5.4 below, $r^2 = x^2 + y^2$, not $x^2 + y^2 + z^2$.]

Problems

5.1 Find the Laplacian of a scalar field $f(r)$ in 2-dimensions. Hence solve Laplace's equation for a circularly symmetric field $f(r)$ in 2-dimensions.

5.2 Show that the potential of a uniform sphere of mass M and radius a is given by

$$\varphi = M/r, \quad r > a ,$$
$$\varphi = M(3a^2 - r^2)/2a^3, \; r < a .$$

[The student may wish to consult references for this problem.]

5.3 Show that the following scalar fields satisfy Laplace's equation

(a) $\varphi = \arctan \frac{y}{x}$,

(b) $\varphi = z \arctan \frac{y}{x}$,

(c) $\varphi = \frac{1}{\sqrt{x^2+y^2+z^2}} \arctan \frac{y}{x}$.

[Hint: use cylindrical polar coordinates.]

5.4 In Chapter 4 we studied functions of the form $z^2 + 1$, e^z, $\cos z$, etc., where $z = x + iy$. If these functions are written in the form $u(x,y) + iv(x,y)$ then, as will be shown in a more advanced course, $u(x,y)$ and $v(x,y)$ are solutions of Laplace's equation. For example

$$z^2+1 = (x+iy)^2 + 1 = (x^2-y^2+1) + i(2xy)$$

$$e^z = e^{x+iy} = e^x \cos y + i\, e^x \sin y$$

$$\cos z = \cos(x+iy) = \cos x \cos iy + \sin x \sin iy$$

$$= \cos x \cosh y + i \sin x \sinh y.$$

(a) Verify by direct calculation that the functions (x^2-y^2+1), $2xy$, $e^x \cos y$, $e^x \sin y$, $\cos x \cosh y$, $\sin x \sinh y$ each satisfy Laplace's equation.

(b) From the invariance of the Laplacian it follows that if we make a change of variables, say

$$x = r\cos\theta\ ,$$
$$y = r\sin\theta\ ,$$

the resulting functions of r,θ will again satisfy Laplace's equation. Apply this change of variables to the first three functions given in (a) and verify that they satisfy Laplace's equation in plane polar coordinates ((5.17) with the term $\frac{\partial^2\varphi}{\partial z^2}$ deleted).

5.5 For what values of n does r^n satisfy Laplace's equation (for $r \neq 0$)? (Use (5.16) and verify the answer to Problem 3.9.)

5.6 (Refer to Problem 8.7 of Chapter 5.) Let C be a 3x3 orthogonal matrix, with $\det C = +1$.

(a) Show that $\lambda = 1$ is an eigenvalue of C.

[Hint: Any equation of the form $\lambda^3+a_1\lambda^2+a_2\lambda-1 = 0$ has a positive root.]

(b) Let $\vec{v}_1$ be a unit eigenvector of C associated with the eigenvalue $\lambda = 1$, and let $\vec{v}_2$ and $\vec{v}_3$ be orthogonal unit vectors, each orthogonal to $\vec{v}_1$. Show that if $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is taken as a new basis C becomes

$$C' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix},$$

where C' is also orthogonal with $\det C' = +1$.

(c) Apply Example 8.2 of Chapter 2 to show that C', and hence also C, represents a rotation about $\vec{v}_1$ as axis.

(d) Find the axis and the angle of the rotation with the matrix

$$\begin{pmatrix} 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \\ 1/3 & -2/3 & 2/3 \end{pmatrix} .$$

Answer. Axis: $(1,0,1)$; angle: $\sin\theta = \frac{\sqrt{8}}{3}$, $\cos\theta = 1/3$.

CHAPTER 8

Boundary Value Problems and Fourier Series

1. Steady State Heat Distribution in a Rectangular Plate.

In his classic treatise Theorie Analytique de la Chaleur (1822) J. B. J. Fourier not only presented us with the tools for finding the flow of heat in solids, but also developed certain mathematical ideas which were received with considerable incredulity at that time.* We shall meet these ideas (Fourier Series) first in the course of determining the steady state temperature distributions in a flat plate.

In Example 3.4 of Chapter 7 we have seen that the steady state temperature distribution in a solid satisfies Laplace's equation $\nabla^2 u = 0$. Consider the problems of determining the steady state temperature distribution in the rectangular plate shown in Figure 1.1. It is assumed that the top and bottom

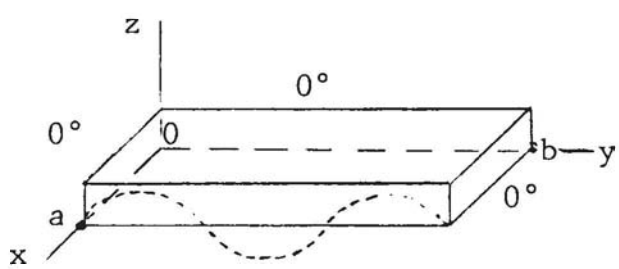


Figure 1.1

*Actually these ideas were already used a century earlier by Daniel Bernoulli and Leonhard Euler in analyzing the vibrations of strings (See Problem 5.10). They were not put on a firm mathematical basis until 1837 by Lejeune-Dirichlet. This work and that of Weierstrass, Riemann and Cantor which followed, initiated the modern developments in mathematical analysis.

surfaces are insulated and that there is no temperature variation in the z-direction. Of the four remaining faces, three are kept at temperature 0° while the fourth is kept at temperature 70 sin $3\pi y/b$, as is indicated schematically by the dotted curve. Since there is no variation in the z-direction we can take the temperature to be a function of x and y only, $u(x,y)$. This function satisfies the following conditions

(1.1) $$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

(1.2) $$u(x,0) = 0$$

(1.3) $$u(x,b) = 0$$

(1.4) $$u(0,y) = 0$$

(1.5) $$u(a,y) = 70 \sin 3\pi y/b \ .$$

It can be proved that the mathematical problem (1.1)-(1.5) has one and only one solution. We leave the proof of this to the reader (See Problem 1.6). We shall proceed on the assumption that there is a unique solution, and attempt to find it.

[Existence and uniqueness proofs are important for the general theory of such systems of equations. The reader who thinks that a mathematical proof is not necessary, on the grounds that the <u>physical problem</u> has a unique solution, is deluding himself. For the mathematical formulation (1.1)-(1.5) is based on various approximations to the physical situation, and the induced inaccuracies may or may not be sufficient to impair the unique solvability. Furthermore, any physical situation is infinitely rich in detail and it is by no means obvious that any set of mathematical conditions is sufficient to characterize it completely.]

Although we have already met many solutions to Laplace's equation (1.1) (cf. Problems 5.3, 5.4 of Chapter 7) none of them satisfy the remaining conditions (1.2)-(1.5). Thus the problem is to find not merely a solution of the differential equation (1.1), but one which satisfies the remaining conditions (1.2)-(1.5) describing the behavior of u at the boundary of the plate. Thus this is called a <u>boundary value problem</u>.

We shall attempt to find the solution in the form of a product of a function of x by a function of z; i.e., $u(x,y) = X(x)Y(y)$. [Note that not every function of x and y can be written in the form of such a product. Our attempt to find a solution in this form is an over-optimistic guess. We shall see that it succeeds in the particular problem we are now solving but will not succeed in general. We shall then learn how to modify the method for a more general case.] With this assumption (1.1)-(1.5) becomes

$$X''Y + XY'' = 0 \tag{1.6}$$

$$X(x)Y(0) = 0 \tag{1.7}$$

$$X(x)Y(b) = 0 \tag{1.8}$$

(1.9) $$X(0)Y(y) = 0$$

(1.10) $$X(a)Y(y) = 70\sin 3\pi y/b \quad .$$

Since if $X(x) \equiv 0$ or if $Y(y) \equiv 0$ it would be impossible to satisfy (1.10) we can replace (1.7)-(1.9) by (1.12)-(1.14) below and put the problem into the form:

(1.11) $$X''Y + XY'' = 0$$

(1.12) $$Y(0) = 0$$

(1.13) $$Y(b) = 0$$

(1.14) $$X(0) = 0$$

(1.15) $$X(a)Y(y) = 70 \sin 3\pi y/b \ .$$

For values of x and y for which $X \neq 0$ and $Y \neq 0$ we can write (1.11) in the form $\frac{Y''}{Y} = -\frac{X''}{X}$. Note that the left side varies only with x, while the right side varies only with y. Consequently both sides are equal to a constant λ, and (1.11) can be replaced by the two ordinary differential equations,

(1.16) $$Y'' = \lambda Y,$$

(1.17) $$X'' = -\lambda X \ .$$

To determine possible values of λ we start with equation (1.16) and the associated boundary conditions (1.12) and (1.13). We first show that λ cannot be zero or positive.

If λ is positive, say $\lambda = k^2$, the general solution of (1.16) is $Y = A \cosh ky + B \sinh ky$. From (1.12), (1.13) we must have

$$0 = A$$

$$0 = A \cosh kb + B \sinh kb \ .$$

Hence $A = B = 0$ and $Y \equiv 0$. Thus we will not be able to satisfy (1.15). Therefore λ is not a positive number.

If we try $\lambda = 0$ in (1.16) we get $Y'' = 0$, which has the general solution $Y = Ay + B$.From (1.12), (1.13) we get

$$0 = B$$

$$0 = Ab + B.$$

Hence again $A = B = 0$ and so $Y \equiv 0$. Thus we cannot satisfy (1.15). Therefore λ is not zero.

Finally let us try negative values for λ, say $\lambda = -k^2$. Then (1.16) becomes $Y'' = -k^2 Y$. The general solution of this differential equation is

$$Y = A \sin ky + B \cos ky \ .$$

From (1.12) we get $B = 0$ so that $Y = A \sin ky$. To satisfy (1.13) we must have

$$0 = A \sin kb \ . \tag{1.18}$$

We cannot choose $A = 0$ or we will again find $Y \equiv 0$ and so will not be able to satisfy (1.15). We can still satisfy condition (1.18) with $A \neq 0$ by selecting $k = \frac{n\pi}{b}$, where n is any integer. Hence $Y(y) = A \sin \frac{n\pi y}{b}$.

Thus we see that the system

$$Y'' = \lambda Y, \quad Y(0) = Y(b) = 0, \tag{1.19}$$

has as its only non-zero solutions the functions

$$Y(y) = A \sin \frac{n\pi}{b} y,$$

corresponding to $\lambda = -\frac{n^2\pi^2}{b^2}$. The reader may already have realized that equations (1.19) specify an eigenvalue problem whose solutions are the eigenvalues $-\frac{n^2\pi^2}{b^2}$ with associated eigenvectors $A \sin \frac{n\pi}{b} y$. This problem is, in fact, essentially the one solved in Example 11.1 of Chapter 5.

We now turn to the equations in x, namely equations (1.17) and (1.14)

$$X'' = -\lambda X = \frac{n^2\pi^2}{b^2} X, \quad X(0) = 0 .$$

The solution

$$X = C \sinh \frac{n\pi}{b} x + D \cosh \frac{n\pi}{b} x$$

that satisfies $X(0) = 0$, is

$$X = C \sinh \frac{n\pi}{b} x.$$

Therefore

$$u(x,y) = X(x)Y(y) = E \sinh \frac{n\pi x}{b} \sin \frac{n\pi y}{b} , \tag{1.20}$$

where we have replaced the arbitrary constant AC by E. Equation (1.20) gives a function $u(x,y)$ which, for any choice of the constant E, and any integer n, will satisfy equations (1.11)-(1.14). What about equation (1.15)? Using (1.20) in (1.15) we get

$$E \sinh \frac{n\pi a}{b} \sin \frac{n\pi y}{b} = 70 \sin 3\pi y/b \quad . \tag{1.21}$$

We can satisfy (1.21) if we choose $n = 3$ and $E = \dfrac{70}{\sinh \frac{3\pi a}{b}}$.

Thus the solution to this problem is

$$u(x,y) = \frac{70}{\sinh \frac{3\pi a}{b}} \sinh \frac{3\pi x}{b} \sin \frac{3\pi y}{b} \ . \tag{1.22}$$

The reader may verify by direct calculation that (1.22) satisfies all the conditions (1.1)-(1.5).

The reader may have felt that while most of the above procedure was straightforward, there was considerable luck involved in being able to solve (1.21). This is indeed the case. Suppose for our next example we consider the same problem but with equation (1.5) replaced by

$$u(a,y) = 70 \sin 3\pi y/b + 112 \sin 8\pi y/b \ . \tag{1.23}$$

The analysis would proceed in exactly the same way except that when we arrive at the step corresponding to (1.21) we have instead

$$E \sinh \frac{n\pi a}{b} \sin \frac{n\pi y}{b} = 70 \sin 3\pi y/b + 112 \sin 8\pi y/b. \tag{1.24}$$

It is impossible to find values of E and n such that (1.24) is true for all values of y. We get around this difficulty in the following way. Note that equations (1.1)-(1.4) are linear and homogeneous; thus if $u_1(x,y)$ and $u_2(x,y)$ each satisfy these equations then so does their sum $u_1(x,y) + u_2(x,y)$. Hence, instead of (1.20), we can try

(1.25) $u(x,y) = E \sinh \frac{n\pi x}{b} \sin \frac{n\pi y}{b} + F \sinh \frac{m\pi x}{b} \sin \frac{m\pi y}{b}$.

The right hand side of (1.25) satisfies (1.1)-(1.4) for any values of the constants E and F and any integer values assigned to n and m. To satisfy (1.23) we take $n = 3$, $m = 8$, $E = 70/\sinh \frac{3\pi a}{b}$, $F = 112/\sinh \frac{8\pi a}{b}$. Hence the function

$$(1.26) \qquad u(x,y) = \frac{70}{\sinh \frac{3\pi a}{b}} \sinh \frac{3\pi x}{b} \sin \frac{3\pi y}{b} + \frac{112}{\sinh \frac{8\pi a}{b}} \sinh \frac{8\pi x}{b} \sin \frac{8\pi y}{b},$$

satisfies equations (1.1)-(1.4) and (1.23), as the student may verify by direct computation. Thus (1.26) solves this problem.

Similarly if the temperature at the face $x = a$ is given by a sum of the form

$$(1.27) \qquad u(a,y) = \sum_{n=1}^{N} K_n \sin n\pi y/b$$

a similar analysis can be carried out. Instead of (1.25) we get

$$(1.28) \qquad u(x,y) = \sum E_n \sinh \frac{n\pi x}{b} \sin \frac{n\pi y}{b} .$$

By selecting the E_n by the formula

$$(1.29) \qquad E_n = \frac{K_n}{\sinh \frac{n\pi a}{b}} \quad \text{for } n = 1,2,\ldots,N.$$

we arrive at the function

$$(1.30) \qquad u(x,y) = \sum_{n=1}^{N} \frac{K_n}{\sinh \frac{n\pi a}{b}} \sinh \frac{n\pi x}{b} \sin \frac{n\pi y}{b}$$

which satisfies equations (1.1)-(1.4) and (1.27) as the student may easily verify by direct computation.

The temperature distribution on the face $x = a$ given by (1.27) is still rather special. Let us now consider the case when it is given by some general function $f(y)$ defined for $0 \leqq y \leqq b$. Equation (1.5) is then replaced by

$$(1.31) \qquad u(a,y) = f(y)$$

The procedure in this case is very similar to that developed above. We assume that there exists a solution $u(x,y)$ in the form of an infinite series of products

$$(1.32) \qquad u(x,y) = \sum_{n=1}^{\infty} X_n(x)Y_n(y)$$

where each term $X_n(x)Y_n(x)$ satisfies the homogeneous conditions (1.1)-(1.4), and consequently the sum does. (Provided that one is permitted to differentiate the series termwise to verify (1.1)). The only non-homogeneous condition (1.31) becomes

$$(1.33) \qquad \sum_{n=1}^{\infty} X_n(a)Y_n(y) = f(y) .$$

The determination of the terms $X_n(x)Y_n(y)$ proceeds in exactly the same steps as equations (1.6)-(1.20) with (1.20) replaced by

$$X_n(x)Y_n(y) = E_n \sinh \frac{n\pi x}{b} \sin \frac{n\pi y}{b} \tag{1.34}$$

and

$$u(x,y) = \sum_{n=1}^{\infty} X_n(x)Y_n(y) = \sum_{n=1}^{\infty} E_n \sinh \frac{n\pi x}{b} \sin \frac{n\pi y}{a} . \tag{1.35}$$

Finally we must satisfy the condition (1.33) which becomes

$$\sum_{n=1}^{\infty} E_n \sinh \frac{n\pi a}{b} \sin \frac{n\pi y}{b} = f(y). \tag{1.36}$$

Can we choose the constants $E_1, E_2, \ldots, E_n, \ldots$ in such a way that (1.36) is satisfied? The remarkable fact is that this can indeed be done for a very wide class of functions $f(y)$. We shall study the theory of such representations in Sections 2 to 4, but for the present let us assume that (1.36) is true and see how to determine the values of the E_n's.

Let k be a fixed positive integer. Multiply each side of (1.36) by $\sin \frac{k\pi y}{b}$ and integrate on y from 0 to b. Assuming that it is permissable to integrate termwise we get

$$\sum_{n=1}^{\infty} E_n \sinh\frac{n\pi a}{b} \int_0^b \sin \frac{k\pi y}{b} \sin\frac{n\pi y}{b}\, dy = \int_0^b f(y) \sin k\pi y/b \, dy. \tag{1.37}$$

Now

$$\int_0^b \sin \frac{k\pi y}{b} \sin \frac{n\pi y}{b}\, dy = \begin{cases} 0 \text{ if } k \neq n & \text{(See Problem 6.10,} \\ \frac{b}{2} \text{ if } k = n . & \text{Chapter 4.)} \end{cases} \tag{1.38}$$

Consequently only one term in the sum on the left side of (1.37) is not zero and we get

$$(1.39) \qquad E_k \left(\sinh \frac{k\pi a}{b}\right) \frac{b}{2} = \int_0^b f(y) \sin \frac{k\pi y}{b}\, dy.$$

Hence

$$(1.40) \qquad E_k = \frac{b_k}{\sinh \frac{k\pi a}{b}} ,$$

where

$$(1.41) \qquad b_k = \frac{2}{b} \int_0^b f(y) \sin \frac{k\pi y}{b}\, dy.$$

Putting this into (1.35) we obtain the solution

$$(1.42) \qquad u(x,y) = \sum_{n=1}^{\infty} \frac{b_n}{\sinh \frac{n\pi a}{b}} \sinh \frac{n\pi x}{b} \sin \frac{n\pi y}{b}$$

where b_n is given by (1.41).

The derivation of (1.42) was formal, rather than rigorous, since we assumed that we could differentiate and integrate certain series termwise. The proposed solution (1.42) can be studied directly however, and, for a large class of functions f(y), can be shown to be the solution to the problem (1.1)-(1.4) and (1.31). We shall not go into the details of this proof, but content ourselves with the formal derivation of the formulas (1.42) and (1.41) as given above. The method, as we shall see later, has other applications. It is called the method of "separation of variables." The student at this stage is expected to obtain only formal solutions to these problems, since the rigorous proof that the formal solution is the actual solution usually requires more advanced mathematical techniques than the student has had.

Problems

1.1 The text shows how to find the steady state temperature distribution in a rectangular plate in which the temperature on three sides is zero and on the fourth side is a prescribed function. Thus one can find functions $u_1(x,y)$, $u_2(x,y)$, $u_3(x,y)$, $u_4(x,y)$ which respectively solve the following four problems.

Problem 1:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

$u(x,0) = 0$

$u(x,b) = 0$

$u(0,y) = 0$

$u(a,y) = f_1(y)$

Problem 2:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

$u(x,0) = 0$

$u(x,b) = 0$

$u(0,y) = f_2(y)$

$u(a,y) = 0$

Problem 3:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

$u(x,0) = 0$

$u(x,b) = f_3(x)$

$u(0,y) = 0$

$u(a,y) = 0$

Problem 4:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

$u(x,0) = f_4(x)$

$u(x,b) = 0$

$u(0,y) = 0$

$u(a,y) = 0.$

Show that the sum $u(x,y) = u_1(x,y) + u_2(x,y) + u_3(x,y) + u_4(x,y)$ gives the steady state temperature distribution in the plate when the temperatures on the four edges are prescribed functions $f_1(y)$, $f_2(y)$, $f_3(x)$, $f_4(x)$.

1.2 Find the steady state temperature, to the nearest degree, at the center of a square plate if three edges are at 0° and the fourth edge at 100°. Answer: 25°.

1.3 Find the steady state temperature distribution in the rectangular plate shown in Figure 1.2 (cf. Problem 1.1).

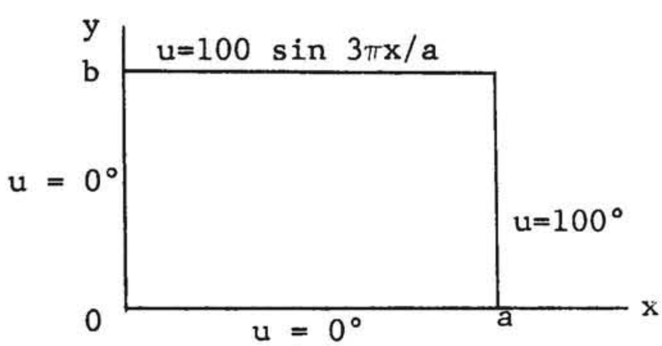


Figure 1.2

Answer:

$$u(x,y) = \frac{100 \sin\frac{3\pi x}{a} \sinh\frac{3\pi y}{a}}{\sinh\frac{3\pi b}{a}} + \frac{400}{\pi} \sum_{n=1,3,5,\cdots} \frac{\sin\frac{n\pi y}{b} \sinh\frac{n\pi x}{b}}{n \sinh\frac{n\pi a}{b}} .$$

1.4 If the temperature variations in the z-direction are taken into account the differential equation for the steady state temperature $u(x,y,z)$ is

(1.43) $$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0.$$

If the temperature is prescribed to be zero on all the faces of a parallelipiped except the face $z = c$ where it is $f(x,y)$ (see Figure 1.3), apply the method in the text to find the temperature $u(x,y,z)$

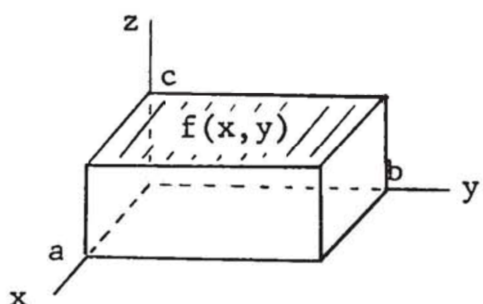


Figure 1.3

[Hint: The boundary conditions are

(1.44) $$u(0,y,z) = 0$$

(1.45) $$u(a,y,z) = 0$$

(1.46) $$u(x,0,z) = 0$$

(1.47) $$u(x,b,z) = 0$$

(1.48) $$u(x,y,0) = 0$$

(1.49) $$u(x,y,c) = f(x,y).$$

If we assume a solution of the form of a sum of products $X(x)Y(y)Z(z)$, where each term $X(x)Y(y)Z(z)$ satisfies (1.43)-(1.48) then show that $\frac{Y''}{Y} + \frac{Z''}{Z} = -\frac{X''}{X} = k^2$ and thus $X = A \sin kx$ where $k = n\pi/a$. Similarly $\frac{Y''}{Y} = k^2 - \frac{Z''}{Z} = -\mu^2$, so that $Y = B \sin \mu y$ where $\mu = m\pi/b$. Hence

$\frac{Z''}{Z} = \left(\frac{n^2}{a^2} + \frac{m^2}{b^2}\right)\pi^2$ and thus show that $Z = C \sinh\sqrt{\frac{n^2}{a^2} + \frac{m^2}{b^2}}\,\pi z$.

From condition (1.49) obtain

$$\sum_m \sum_n C_{mn} \sin\frac{n\pi x}{a} \sin\frac{m\pi y}{b} \sinh\sqrt{\frac{n^2}{a^2} + \frac{m^2}{b^2}}\;\pi c = f(x,y);$$

and then

$$(1.50)\quad C_{mn} = \frac{4}{ab \sinh\sqrt{\left(\frac{n^2}{a^2} + \frac{m^2}{b^2}\right)}\pi c} \int_0^b\!\!\int_0^a f(x,y)\sin\frac{n\pi x}{a}\sin\frac{m\pi x}{b}\,dx\,dy$$

Finally the solution is

$$(1.51)\quad u(x,y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{mn} \sin\frac{n\pi x}{a} \sin\frac{m\pi y}{b} \sinh\sqrt{\frac{n^2}{a^2} + \frac{m^2}{b^2}}\,\pi z$$

where C_{mn} is given by (1.50)]

1.5 Use the results of Problem 1.4 to find the temperature (to one decimal place) of a cubical block all of whose faces except one are at 0° and whose other face is at 100°.

1.6 Prove, by providing the reasons for the following sequence of steps, that the mathematical problem (1.1)-(1.5) has a unique solution.

(a) There exists a solution; in fact (1.22) is a solution, as may be verified easily by direct computation.

(b) If U(x,y) and v(x,y) are two solutions, let w(x,y) = u(x,y) - v(x,y). Then w(x,y) satisfies the conditions

(i) $\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = 0$

(ii) $w(x,0) = 0$

(iii) $w(x,b) = 0$

(iv) $w(0,y) = 0$

(v) $w(a,y) = 0.$

(c) $$0 = \int_0^b \int_0^a w\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right) dx\ dy$$

$$= \int_0^b \int_0^a w \frac{\partial^2 w}{\partial x^2} dx\ dy + \int_0^a \int_0^b w \frac{\partial^2 w}{\partial y^2} dy\ dx$$

$$= \int_0^b \left[w \frac{\partial w}{\partial x} \Big|_{x=0}^{a} - \int_0^a \left(\frac{\partial w}{\partial x}\right)^2 dx \right] dy$$

$$+ \int_0^a \left[w \frac{\partial w}{\partial y} \Big|_{y=0}^{b} - \int_0^b \left(\frac{\partial w}{\partial x}\right)^2 dy \right] dx$$

$$= -\int_0^b \int_0^a \left[\left(\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2 \right] dx\ dy.$$

Hence $\frac{\partial w}{\partial x} = 0$ and $\frac{\partial w}{\partial y} = 0$. Hence w(x,y) is a constant.

Hence $w(x,y) \equiv 0$. Hence $u(x,y) \equiv v(x,y)$.

2. Fourier Series on the Interval $(-\pi,\pi)$

A series of the form

$$(2.1)\quad A_0 + A_1 \cos x + B_1 \sin x + A_2 \cos 2x + B_2 \sin 2x + \dots + A_n \cos nx + B_n \sin nx + \dots$$

where the A_n and B_n are constants, is called a trigonometric series. More briefly we write it in the form

$$(2.2)\qquad \sum_{n=0}^{\infty}(A_n \cos nx + B_n \sin nx)$$

the term B_0 being irrelevant since $\sin 0x = 0$.

For those values of x for which the series (2.2) converges we can define a function f(x)

$$(2.3)\qquad f(x) = \sum_{n=0}^{\infty}(A_n \cos nx + B_n \sin nx).$$

Now, just as we passed from Power Series to MacLaurin Series, we shall pass from Trigonometric Series to Fourier Series. That is, we now shall suppose that a certain function f(x) has been given and inquire whether it can indeed be represented in the form (2.3). If f(x) can be represented in the form (2.3), what are the values of the constants A_n, B_n? We shall proceed first in a formal manner. By direct integration the student may verify the following facts. If m and n are distinct non-negative integers, we have

(2.4) $$\int_{-\pi}^{\pi} \cos mx \cos nx \, dx = 0$$

(2.5) $$\int_{-\pi}^{\pi} \cos mx \sin nx \, dx = 0$$

(2.6) $$\int_{-\pi}^{\pi} \sin mx \sin nx \, dx = 0 \ ;$$

If $m = n \neq 0$ we have

(2.7) $$\int_{-\pi}^{\pi} \cos mx \cos nx \, dx = \int_{-\pi}^{\pi} \cos^2 mx \, dx = \pi$$

(2.8) $$\int_{-\pi}^{\pi} \cos mx \sin nx \, dx = \int_{-\pi}^{\pi} \cos mx \sin mx \, dx = 0$$

(2.9) $$\int_{-\pi}^{\pi} \sin mx \sin nx \, dx = \int_{-\pi}^{\pi} \sin^2 mx \, dx = \pi$$

while if $m = n = 0$ we have

(2.10) $$\int_{-\pi}^{\pi} \cos mx \cos nx \, dx = \int_{-\pi}^{\pi} dx = 2\pi.$$

[Note: These facts can be expressed in the language of Chapter 5 as follows. Let $C(-\pi,\pi)$ be the vector space of functions continuous on the interval $(-\pi,\pi)$ with the usual rule for addition and multiplications by scalars. Also, let the inner product on $C(-\pi,\pi)$ be given by $f \cdot g = \int_{-\pi}^{\pi} f(x)g(x)dx$. Then the functions

1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$,... $\cos nx$, $\sin nx$,...

are pairwise orthogonal vectors in $C(-\pi,\pi)$. Furthermore

$$\|\cos nx\| = \sqrt{\pi}, \ \|\sin nx\| = \sqrt{\pi}, \quad \|1\| = \sqrt{2\pi}.$$

See Problem 11.2 of Chapter 5.]

Let m be any fixed positive integer. Multiply each side of equation (2.3) by cos mx and integrate from $-\pi$ to π. Assuming that termwise integration is permissible we obtain

$$\int_{-\pi}^{\pi} f(x) \cos mx \, dx = \sum_{n=0}^{\infty} (A_n \int_{-\pi}^{\pi} \cos nx \cos mx \, dx$$

$$+ B_n \int_{-\pi}^{\pi} \sin nx \cos mx \, dx)$$

We see from formulas (2.4), (2.5), (2.7), (2.8) that each integral on the right side will be zero except the term in which n=m, where we have $A_m \int_{-\pi}^{\pi} \cos mx \cos mx \, dx = \pi A_m$. Hence we find

$$(2.11) \qquad A_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx \, dx, \ (m \geqq 1).$$

Similarly by multiplying each side of (2.3) by sin mx and integrating formally termwise we find

$$(2.12) \qquad B_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx \, dx, \ (m \geqq 1)$$

Finally to evaluate A_0 we multiply each side of (2.3) by cos 0X ($\equiv 1$) and integrate termwise to obtain, using Equation (2.10)

$$(2.13) \qquad A_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, dx.$$

Since the formula (2.13) has a factor $\frac{1}{2}$ which formulas (2.11) and (2.12) do not have, it is conventional to replace the A_n, B_n by a_n, b_n where $a_0 = 2A_0$ and $a_n = A_n$, $b_n = B_n$ for $n \geqq 1$, as we do below. The student can always avoid confusion by remembering simply that the constant term in the expansion (2.3) is always the average of the function f(x) over the interval.

Now, regardless of whether or not these manipulations of the series (2.2) are justifiable, we can <u>define</u> the Fourier Series of the function f(x) by (2.2) where the constants A,B are given by equations (2.11), (2.12), (2.13). Let us state this explicitly using the conventional notation.

<u>Definition 2.1</u>. Let f(x) be a given function, such that the integrals in equations (2.15), (2.16), below exist. The <u>Fourier Series for f(x)</u> is the series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) , \tag{2.14}$$

where the constants a_n, b_n, called the "Fourier coefficients," are given by

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \tag{2.15}$$

and

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx. \tag{2.16}$$

[Note: In the language of Chapter 5, again, the orthogonal set we are now dealing with is

$$\frac{1}{2},\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x\ ,\ldots,\ \cos nx,\ \sin nx\ ,\ldots$$

The Fourier Coefficients, given by equations (2.15) and (2.16) can be written as

$$a_n = \frac{f\cdot(\cos nx)}{\|\cos nx\|^2}\ ,\ b_n = \frac{f\cdot(\sin nx)}{\|\sin nx\|^2}\ ,\ n = 1,2,\ldots$$

$$a_0 = \frac{f\cdot(\frac{1}{2})}{\|\frac{1}{2}\|^2}\ .$$

If we consider the finite set of (2N+1) functions:

$$\{\frac{1}{2},\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ldots,\ \cos Nx,\ \sin Nx\}$$

and let W be the vector subspace spanned by these functions (i.e., the set of all linear combinations of these functions) then Theorem 3.4 of Chapter 5 tells us that any function f in $C(-\pi,\pi)$ can be expressed in the form

$$f = w + u$$

where w is the sum of the first (2N+1) terms of the Fourier series for f, i.e., $\frac{a_0}{2} + \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx)$, where a_n, b_n are given by (2.15), (2.16) and u is orthogonal to each of the (2N+1) functions

$$\{\tfrac{1}{2}, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos Nx, \sin Nx\}$$

Furthermore Theorem 3.5 of Chapter 5 tells us that among all the possible linear combinations of these functions:

$$g(x) = \frac{\alpha_0}{2} + \sum_{n=1}^{N} (\alpha_n \cos nx + \beta_n \sin nx)$$

the <u>mean</u> <u>squared</u> <u>error</u>,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x) - g(x)]^2 \, dx ,$$

is least when the α_n's and β_n's are taken to be equal to the Fourier Coefficients a_n, b_n.

Finally from the result of Problem 3.7 of Chapter 5 we get

$$\sum_{i=1}^{m} \frac{(v \cdot v_i)^2}{\|v_i\|^2} = \|v\|^2 - \|u\|^2 \leqq \|v\|^2 \text{ or } \sum_{i=1}^{m} \left(\frac{v \cdot v_i}{\|v_i\|^2}\right)^2 \|v_i\|^2 \leqq \|v\|^2$$

which means that

$$\frac{\pi}{2} a_0^2 + \pi(a_1^2 + b_1^2 + \ldots + a_N^2 + b_N^2) \leqq \int_{-\pi}^{\pi} f^2(x) \, dx.$$

From the Axiom of Continuity (Chapter 3, Section 11, Property F) it follows that the sum of the squares of the Fourier Coefficients converges and that

$$\frac{1}{2} a_0^2 + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \leqq \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) \, dx.$$

See Problem 10.5 of Chapter 5.]

Now the question is, "Does the Fourier Series for $f(x)$ converge to $f(x)$ at each point x in the interval $(-\pi \leqq x \leqq \pi)$?" Theorem 2.1 below will deal with this question. To set the stage for that we need to make one remark and define some terms.

Definition 2.2. A function $g(x)$ is said to be periodic of period T if $g(x+T) \equiv g(x)$.

Remark: Since each term on the right side of equation (2.14) is periodic of period 2π, it is clear that if the Fourier series does converge to $f(x)$ then $f(x)$ is periodic of period 2π. Thus in Theorem 2.1 we shall postulate that $f(x)$ is periodic of period 2π. After that we shall see how the theorem can still be used when $f(x)$ is not periodic.

Notation: The limit of $f(x)$ as x approaches a number a through values larger than a is denoted by $\lim_{x \to a+} f(x) = f(a+)$. To be precise $f(a+) = b$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that $|f(x) - b| < \epsilon$ whenever $a < x < a + \delta$. Similarly the limit of $f(x)$ as x approaches a through values less than a is denoted by $\lim_{x \to a-} (x) = f(a-)$.

Definition 2.3. A function $f(x)$ is sectionally smooth on an interval (a,b) if the interval can be partioned into a finite number of sub-intervals:

$$(a,x_1),\ (x_1,x_2),\ldots,(x_n,b) \qquad \text{(See Figure 2.1)}$$

such that the following 3 conditions are satisfied.

1. In the interior of each subinterval f(x) is differentiable.
2. As x approaches an endpoint of any subinterval through values in the subinterval f(x) approaches a limit. That is, the limits $f(a+)$, $f(x_1-)$, $f(x_1+),\ldots,f(x_n-)$, $f(x_n+)$, $f(b-)$ exist.
3. The "derivatives" exist at the endpoints in the sense that the limits

$$\lim_{x \to a+} \frac{f(x)-f(a+)}{x-a}, \quad \lim_{x \to x_1-} \frac{f(x_1-)-f(x)}{x_1-x},$$

$$\ldots, \quad \lim_{x \to b-} \frac{f(b-)-f(x)}{b-x} \text{ exist.}$$

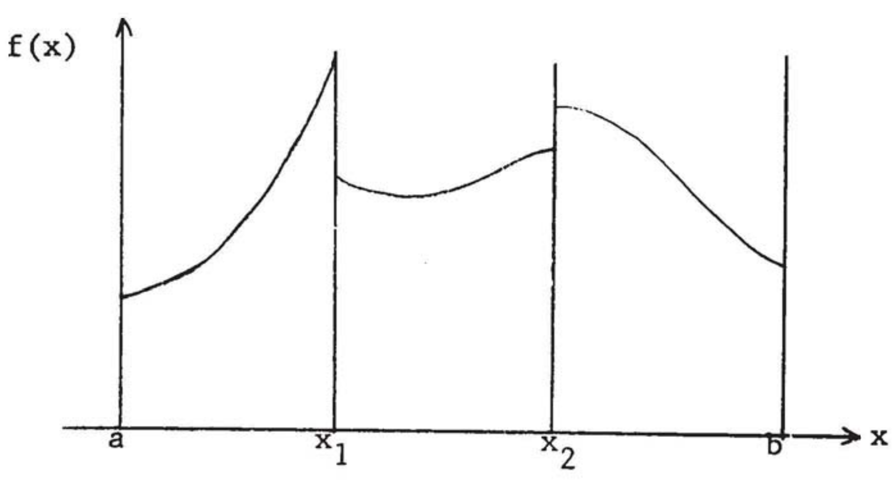


Figure 2.1

<u>Theorem 2.1</u>. If $f(x)$ is periodic of period 2π and is sectionally smooth in the interval $(-\pi,\pi)$ then the Fourier series for $f(x)$ (equations (2.14) - (2.16)) converges to $f(x)$ at all points x at which $f(x)$ is continuous and converges to the number $\frac{f(x+) + f(x-)}{2}$ (i.e., the <u>average</u> of the left- and right-hand limits) at points x at which $f(x)$ is discontinuous.

Note. Theorem 2.1 can be stated more simply as follows. "If $f(x)$ is periodic of period 2π and is sectionally smooth in the interval $(-\pi,\pi)$ then the Fourier Series for $f(x)$ converges to $\frac{f(x+) + f(x-)}{2}$.)

For a proof of this theorem see Sokolnikoff, I. S. and Redheffer, R. M., <u>Mathematics of Physics and Modern Engineering</u>, McGraw-Hill Book Co., New York, 1958, p. 204. We shall not give a proof here.

Suppose we have a function $f(x)$ which is sectionally smooth in the interval $(-\pi,\pi)$ but not periodic of period 2π. Its Fourier Series is again given by equations (2.14)-(2.16), but of course the Fourier Series does not in general represent $f(x)$ outside the interval $(-\pi,\pi)$. If however we are interested in the function $f(x)$ only for x in the interval $(-\pi,\pi)$, then the Fourier Series might represent it in that interval. To enable ourselves to apply Theorem 2.1, we introduce a new function, called the <u>periodic extension of $f(x)$ from the interval $(-\pi,\pi)$</u>. This is the function $g(x)$ defined by

$$g(x) = f(x), \quad -\pi \leqq x < \pi$$

$$g(x) = g(x + 2\pi) \quad \text{all } x \qquad \text{(See Figure 2.2).}$$

The function $g(x)$ coincides with $f(x)$ in the interval $-\pi \leqq x < \pi$. Hence the Fourier Series for $g(x)$ is also given by Equations (2.14)-(2.16). Theorem 2.1 may now be applied to $g(x)$. The fact that the Fourier Series converges to $\frac{g(x+) + g(x-)}{2}$ for all x implies that for all x such that $-\pi < x < \pi$ the Fourier Series will converge to $\frac{f(x+) + f(x-)}{2}$, while for $x = \pm\pi$ it will converge to $\frac{f(-\pi+) + f(\pi-)}{2}$.

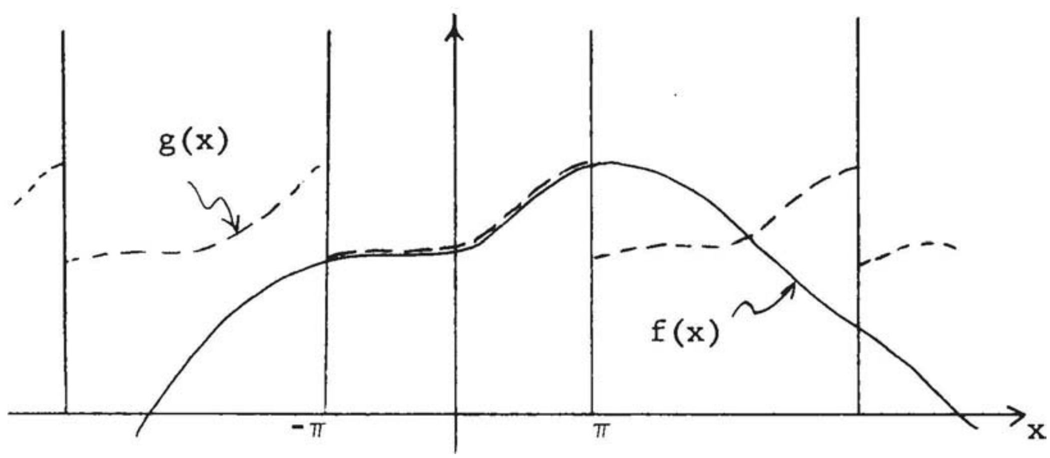


Figure 2.2

Example 2.1. Let $f(x) = x$. The periodic extension of $f(x)$ from the interval $(-\pi,\pi)$ is given by

$$g(x) = x \qquad -\pi \leqq x < \pi$$

$$g(x + 2\pi) = g(x), \quad \text{all } x.$$

Both $f(x)$ and $g(x)$ are sketched in Figure 2.3.

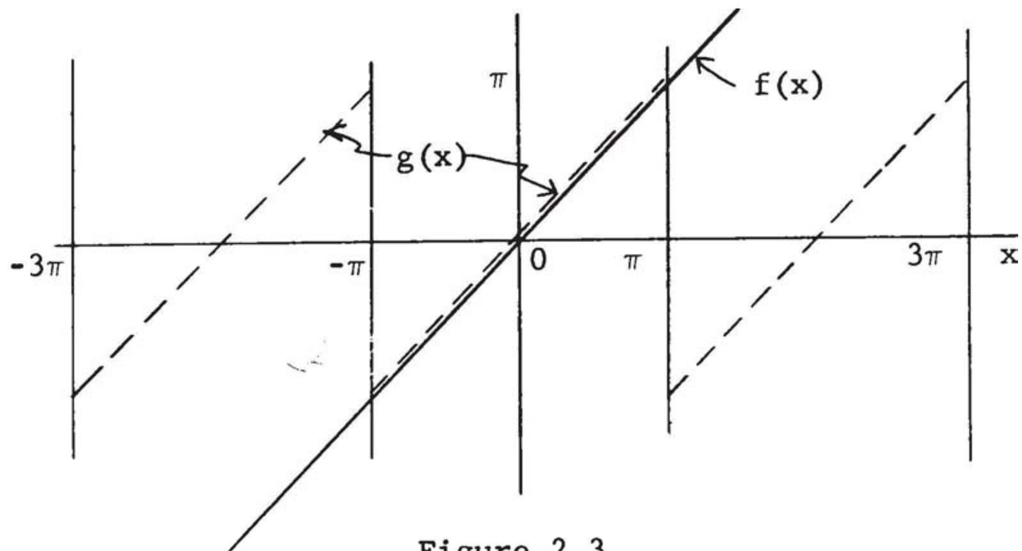


Figure 2.3

The Convergence Theorem (2.1) is concerned with the function $g(x)$. The Fourier Series for $f(x)$ and $g(x)$ are the same and are given by

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos nx + b_n \sin nx\right),$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx \, dx = \frac{1}{\pi}\int_{-\pi}^{\pi} x \cos nx \, dx = 0 \ ,$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin nx \, dx = \frac{1}{\pi}\int_{-\pi}^{\pi} x \sin nx \, dx$$

$$= \frac{1}{\pi}\left[-\frac{x \cos nx}{n} + \frac{\sin nx}{n^2}\right]_{-\pi}^{\pi} = -\frac{2 \cos n\pi}{n}$$

$$= \frac{-2(-1)^n}{n} = \frac{2(-1)^{n+1}}{n} \ .$$

Thus the Fourier Series for f(x) or g(x) is given by

$$(2.17) \qquad 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}\sin nx}{n} = 2\left(\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \dots + \frac{(-1)^{n+1}\sin nx}{n} + \dots\right).$$

The convergence theorem, (Theorem 2.1) then asserts that this series converges to $\frac{g(x+) + g(x-)}{2}$. This function is sketched in Figure 2.4 along with the first three partial sums:

$$S_1(x) = 2\sin x,$$

$$S_2(x) = 2\left(\sin x - \frac{\sin 2x}{2}\right),$$

$$S_3(x) = 2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3}\right).$$

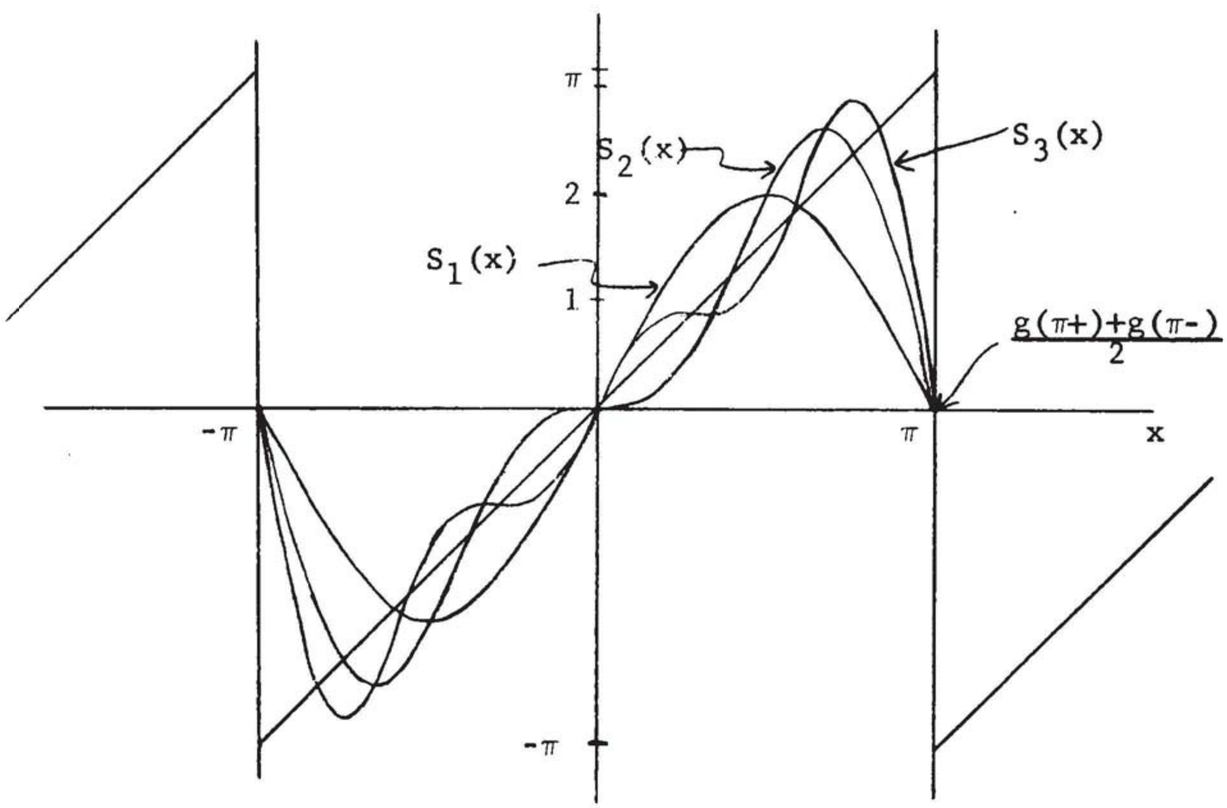


Figure 2.4

Note in particular that at $x = 0, \pm\pi, \pm 2\pi \cdots$ the value of $\frac{g(x+) + g(x-)}{2}$ is zero and the Fourier Series indeed converges to zero.

Although it is slightly extraneous to our present purposes, one may use Fourier Series to obtain formulas for the sum of certain series. Thus for example we see from Figure 2.4 that at $x = \pi/2$, $g(x) = \frac{\pi}{2}$. Therefore in virtue of Theorem 2.1 we have

$$\frac{\pi}{2} = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1} \sin \frac{n\pi}{2}}{n} = 2 \left(\sin \frac{\pi}{2} - \frac{\sin \pi}{2} + \frac{\sin \frac{3\pi}{2}}{3} - \ldots \right)$$

$$= 2 \left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \ldots\right).$$

Thus we have the summation formula

$$(2.18) \qquad 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4} .$$

Example 2.2. Figure 2.5 shows a sketch of the function

$$f(x) = -1 , \; x \leq 0 ,$$
$$f(x) = x , \; x > 0.$$

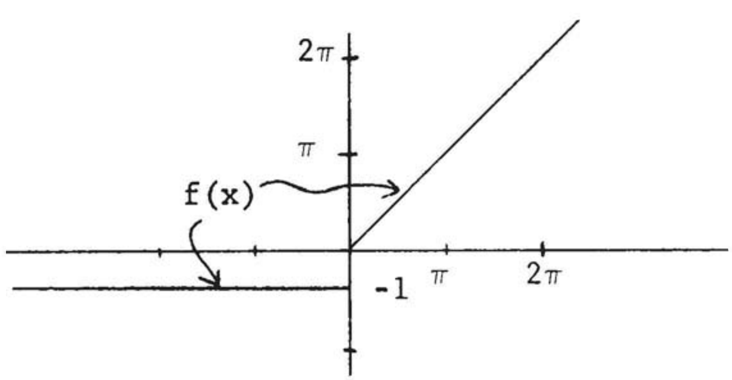


Figure 2.5

Its periodic extension from the interval $(-\pi,\pi)$ is given by

$$g(x) = -1 \ , \ -\pi \leqq x \leqq 0 \ ,$$
$$g(x) = \ x \ , \ \ 0 < x < \pi \ ,$$
$$g(x + 2\pi) = g(x) \qquad \text{all } x,$$

as is sketched in Figure 2.6.

Note that this function is described by a different algebraic expression in each of the two intervals $-\pi \leqq x \leqq 0$ and $0 < x < \pi$. When integrating such a function from $-\pi$ to $+\pi$ we break the integral into two parts; one going from $-\pi$ to 0, the other from 0 to π. In the first we set $g(x) = -1$, and in the second we set $g(x) = x$.

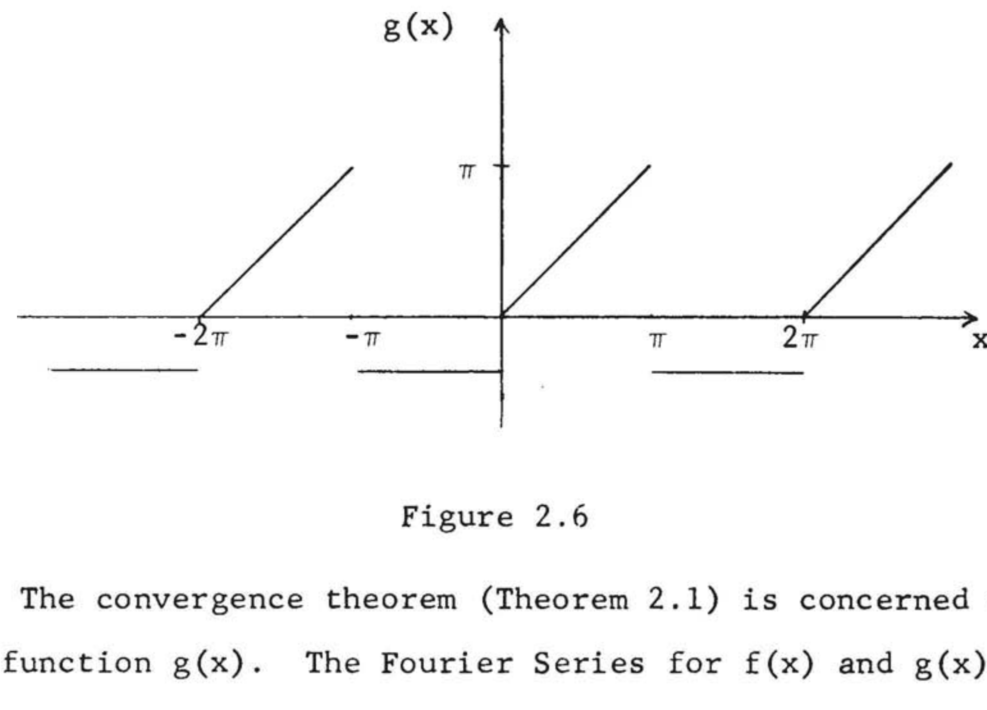


Figure 2.6

The convergence theorem (Theorem 2.1) is concerned with the function $g(x)$. The Fourier Series for $f(x)$ and $g(x)$ are the same and are given by

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

where

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx = \frac{1}{\pi} \int_{-\pi}^{0} (-\cos nx)dx + \frac{1}{\pi} \int_{0}^{\pi} x \cos nx \, dx$$

$$= \frac{-\sin nx}{\pi n} \Bigg|_{-\pi}^{0} + \frac{1}{\pi} \left(\frac{x \sin nx}{n} + \frac{\cos nx}{n^2}\right) \Bigg|_{0}^{\pi}$$

$$= \frac{1}{\pi n^2} (\cos n\pi - 1) = \frac{(-1)^n - 1}{\pi n^2}, \text{ for } n \neq 0 \ ;$$

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)dx = \frac{1}{\pi} \int_{-\pi}^{0} (-1)dx + \frac{1}{\pi} \int_{0}^{\pi} x \, dx = -1 + \frac{1}{\pi} \frac{x^2}{2} \Bigg|_{0}^{\pi}$$

$$= \frac{\pi}{2} - 1;$$

and

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx = \frac{1}{\pi} \int_{-\pi}^{0} (-\sin nx)dx + \frac{1}{\pi} \int_{0}^{\pi} x \sin nx \, dx$$

$$= \frac{1}{\pi} \frac{\cos nx}{n} \Bigg|_{-\pi}^{0} + \frac{1}{\pi} \left(\frac{-x \cos nx}{n} + \frac{\sin nx}{n^2} \right) \Bigg|_{0}^{\pi}$$

$$= \frac{1 - \cos n\pi}{\pi n} - \frac{\cos n\pi}{n} = \frac{1 - (1 + \pi) \cos n\pi}{n\pi} = \frac{1 - (1 + \pi)(-1)^n}{n\pi}.$$

Hence

(2.19) $\frac{g(x+) + g(x-)}{2}$

$$= \frac{1}{2}\left(\frac{\pi}{2} - 1\right) + \frac{1}{\pi}\sum_{n=1}^{\infty}\left(\frac{(-1)^n-1}{n^2}\cos nx + \frac{1-(1+\pi)(-1)^n}{n}\sin nx\right).$$

For example at $x = 0$, $\frac{g(x+) + g(x-)}{2} = -\frac{1}{2}$ (from Figure 2.6).

Hence from (2.19) we get

$$-\frac{1}{2} = \frac{1}{2}\left(\frac{\pi}{2} - 1\right) + \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n-1}{n^2}. \tag{2.20}$$

The numerator in the series is zero when n is even, hence we can write (2.20) in the form

$$\frac{\pi^2}{8} = \sum_{n\ \mathrm{odd}}\frac{1}{n^2} = \frac{1}{1} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \ldots = \sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}. \tag{2.21}$$

Similarly at $x = \pi/2$, we get

$$\frac{\pi}{2} = \frac{1}{2}\left(\frac{\pi}{2} - 1\right) + \frac{1}{\pi}\sum_{k=1}^{\infty}\frac{2+\pi}{2k-1}\sin(2k-1)\frac{\pi}{2},$$

or

$$\frac{\pi}{2} = -\frac{1}{2} + \frac{2+\pi}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{2k-1},$$

i.e.,

$$\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{2k-1} = \frac{\pi}{4}, \tag{2.22}$$

as in (2.18).

The sequence of partial sums in equation (2

$$S_0(x) = \frac{\pi}{4} - \frac{1}{2}$$

$$S_1(x) = \frac{\pi}{4} - \frac{1}{2} + \frac{1}{\pi}\,(-2\cos x + (2 + \pi)\sin x)$$

$$S_2(x) = \frac{\pi}{4} - \frac{1}{2} + \frac{1}{\pi}\,(-2\cos x + (2 + \pi)\sin x) - \frac{1}{2}\sin 2x.$$

These functions are graphed in Figure 2.7 along with g(x).

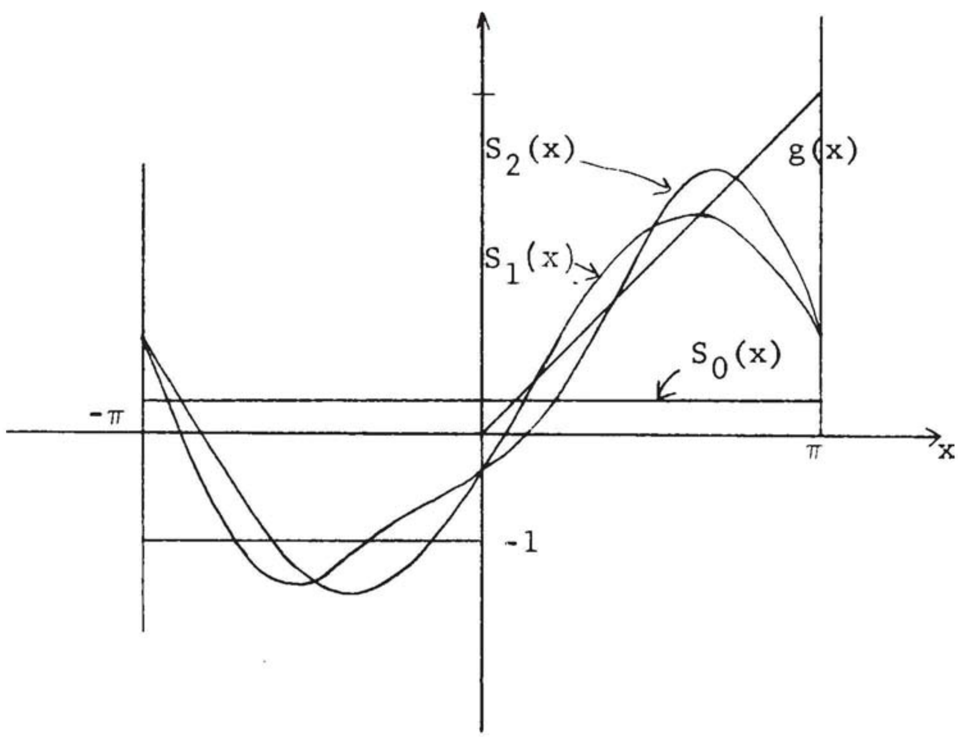


Figure 2.7

R 6 +

Problems

detail the computations of Example 2.1.

he graphs carefully and add the graph of

to your version of Figure 2.4.

[Note: You may find it advisable to have the computer do the arithmetic for you.]

(c) Do the same for Example 2.2 and Figure 2.7.

2.2 Repeat the procedure of Example 2.1 but starting with

(a) the sawtooth wave (see Figure 2.8).

$f(x) = |x|, \quad -\pi \leqq x < \pi,$

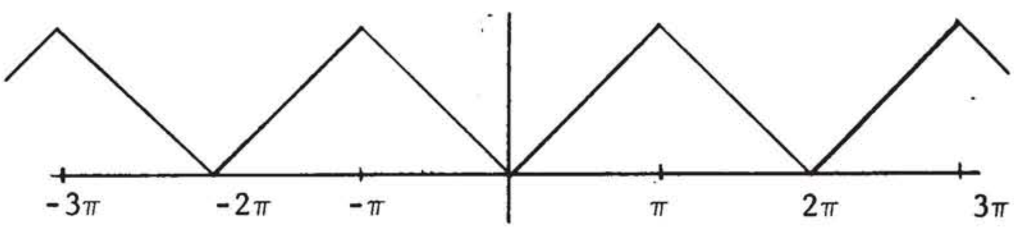


Figure 2.8

Answer: $f(x) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n \text{ odd}} n^{-2} \cos nx.$

(b) the square-sine curve (see Figure 2.9).

$\text{Sin}(x) = -1, \ -\pi \leqq x \leqq 0,$

$\text{Sin}(x) = 1, \ 0 < x < \pi,$

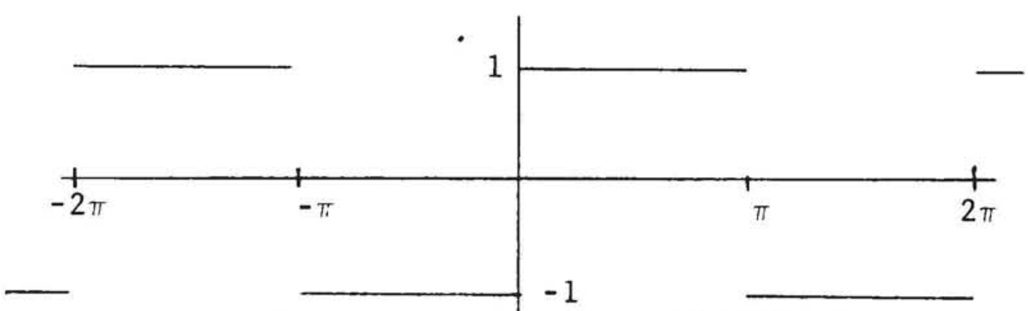


Figure 2.9

(c) the square-cosine curve (see Figure 2.10),

$\text{Cos}\ x = -1\ ,\ -\pi \leqq x \leqq -\pi/2\ ,\ \pi/2 < x < \pi,$

$\text{Cos}\ x = +1\ ,\ -\pi/2 < x \leqq \pi/2\ .$

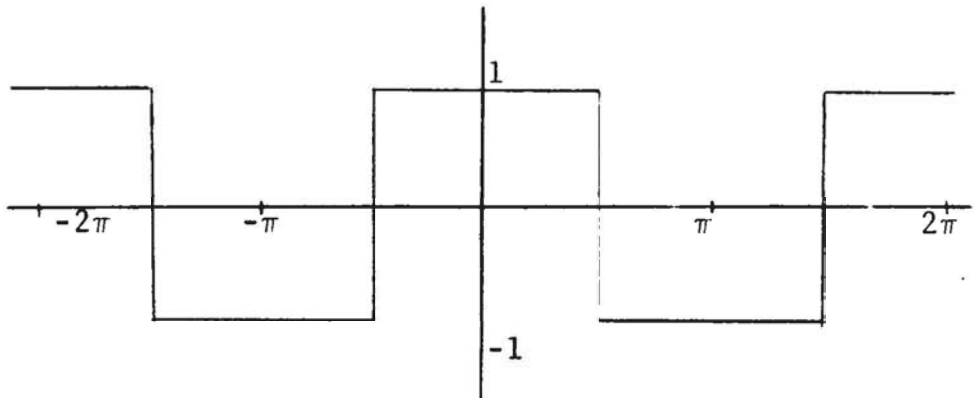


Figure 2.10

(d) the ramp function,

$f(x) = 0\ ,\ x < 0\ ,$

$f(x) = x\ ,\ 0 \leqq x \leqq \pi\ ,$

$f(x) = \pi\ ,\ x > \pi.$

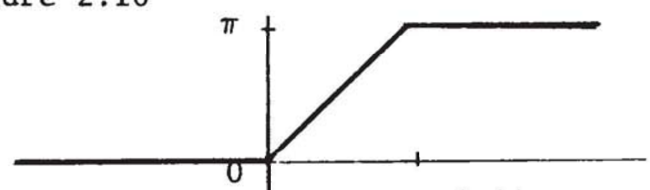


Figure 2.11

(e) the step function,

$f(x) = 0\ ,\ x < 0\ ,$

$f(x) = 1\ ,\ x \geqq 0.$


Figure 2.12

Ans. $\frac{1}{2} + \sum_{m=1}^{\infty} \frac{\sin(2m-1)x}{2m-1}$.

(f) the half-rectified sine wave,

$f(x) = 0\ ,\ -\pi \leqq x \leqq 0\ ,$

$f(x) = \sin x\ ,\ 0 < x < \pi.$

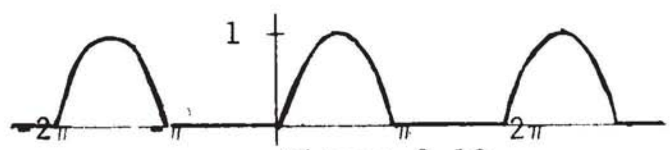


Figure 2.13

Check your Fourier Series by letting $x = \pi/2$ and verifying the formula

$$\frac{\pi}{4} - \frac{1}{2} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{4n^2-1}\ .$$

(g) the full rectified sine wave $f(x) = |\sin x|$.

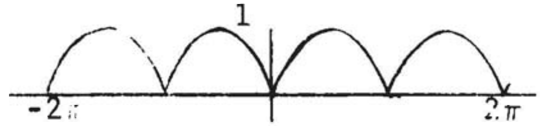


Figure 2.14

(h) The function $f(x) = x^2$. Check your result by verifying equation (2.21) and also the summation formulas:

$$\frac{\pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2},$$

$$\frac{\pi^2}{12} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}.$$

2.3 Show that the Fourier Series for $f(x)$ can be written in the form:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)dt + \frac{1}{\pi}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} f(t)\cos n(x-t)dt.$$

2.4 Suppose that $f(x)$ is given on the interval $(-\pi,\pi)$. Discuss the relative merits of a Taylor Series representation versus a Fourier Series representation (e.g., conditions on $f(x)$ under which the series is defined, convergence to $f(x)$, the goodness of fit at various values of x in the interval, termwise differentiability and integrability, etc.) You may consult other texts.

2.5 Show that the series (2.2) can be written as the real part of the power series $\sum_{n=0}^{\infty} C_n z^n$, where $C_n = A_n - iB_n$ and $z = e^{ix}$.

2.6 In complex form (see Problem 10.5 of Chapter 5) the Fourier Series of $f(x)$ is

$$\sum_{n=-\infty}^{\infty} a_n e^{inx}, \tag{2.23}$$

where

(2.24) $$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-inx} f(x)\,dx.$$

(a) Taking $f(x) = x$, obtain the Fourier Series of the function x in the form

$$\sum \frac{(-1)^{n+1}}{in} e^{inx} ,$$

where $\sum$ goes from $-\infty$ to ∞ but omits $n = 0$. By pairing off the terms with the same value of $|n|$ reduce this sum to (2.17).

(b) Solve Problem 2.2f by this method. [Hint: Express $\sin x$ in terms of exponentials before integrating.]

3. Change of Interval

At the end of the previous section we saw how any function $f(x)$ which is sectionally smooth in the interval $(-\pi,\pi)$ could be represented by its Fourier Series in that interval. Suppose we now have a function $g(x)$ which is sectionally smooth on an interval $|x-a| < r$. (See Figure 3.1). By making a change of variable which maps the interval $|x-a| < r$ into the interval $-\pi < y < \pi$ we can reduce this problem to the previous case. Specifically let $\frac{x-a}{r} = \frac{y}{\pi}$, i.e., let

(3.1) $$y = \frac{\pi}{r}(x-a),\quad x = a + \frac{ry}{\pi},$$

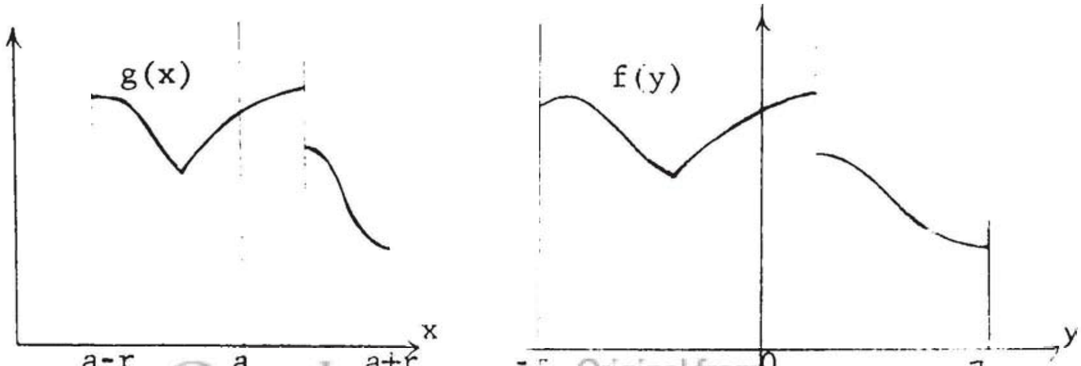


Figure 3.1

and let

$$(3.2) \qquad f(y) = g(x) = g(a + \frac{ry}{\pi}) \ , \ -\pi < y < \pi .$$

Applying Theorem 2.1 to f(y) we have

$$(3.3) \qquad \frac{f(y+) + f(y-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos ny + b_n \sin ny)$$

where

$$(3.4) \qquad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos ny \, dy, \ b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin ny \, dy.$$

Using (3.1) and (3.2) we obtain from (3.3) and (3.4) respectively

$$(3.5) \qquad \frac{g(x+) + g(x-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos \frac{n\pi}{r}(x-a) + b_n \sin \frac{n\pi}{r}(x-a))$$

where

$$(3.6) \qquad a_n = \frac{1}{r} \int_{a-r}^{a+r} g(x) \cos \frac{n\pi}{r}(x-a) \, dx$$

and

$$(3.7) \qquad b_n = \frac{1}{r} \int_{a-r}^{a+r} g(x) \sin \frac{n\pi}{r}(x-a) \, dx.$$

The series on the right of equation (3.5) with the coefficients given by equations (3.6) and (3.7) is called the <u>Fourier Series for g(x) on the interval</u> $|x-a| < r$. Equation (3.5) gives a valid representation when g(x) is sectionally smooth on the interval $|x-a| < r$.

Problems

3.1 Find the Fourier Series for the indicated functions on the indicated interval, and sketch (for comparison) the graph of the function and the first three partial sums of its Fourier Series on that interval.

(a) $g(x) = x, \ |x| \leqq 1$

Answer: $\sum_{n=1}^{\infty} \frac{2}{n\pi} (-1)^{n+1} \sin n\pi x$

(b) $g(x) = \sin x, \ |x| \leqq 1$

Answer: $\sum_{n=1}^{\infty} \frac{1.68 n\pi (-1)^{n+1}}{(n\pi)^2 - 1} \sin n\pi x$, $(2 \sin 1 \cong 1.68)$

(c) $g(x) = e^x, \ |x-1| \leqq 3$

(d) $f(x) = \begin{cases} 1 & , \ 0 \leqq x < 1/2 \\ 0 & , \ 1/2 \leqq x \leqq 1 \end{cases}$ (Interval $0 \leqq x \leqq 1$)

Answer: $\frac{1}{2} + \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{\sin(4n+2)\pi x}{2n+1}$.

3.2 (a) Find the Fourier Series for the function shown in Figure 3.2 on the interval $0 \leqq x < 2$.

(b) Sketch the function to which the series in (a) converges in the interval $-2 \leqq x \leqq 4$ indicating in particular the value at any point of discontinuity.

Answer: $\frac{1}{4} - \frac{2}{\pi^2} \sum_{n \text{ odd}} \frac{\cos n\pi(x-1)}{n^2} + \frac{1}{\pi} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sin n\pi(x-1)}{n}$.

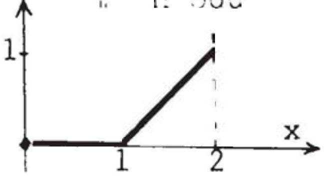


Figure 3.2

3.3 (a) Show that the complex form of the Fourier Series (Problem 2.6), for the interval $a < x < b$, is

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{incx}$$

with

$$a_n = \frac{1}{b-a}\int_a^b e^{-incx} f(x)\,dx \ ,$$

where

$$c = 2\pi/(b-a).$$

(b) Use this formulation to show that

$$\frac{\cosh x}{\sinh 1} = 1 + 2\sum_{n=1}^{\infty} \frac{(-1)^n}{1+n^2\pi^2}\cos n\pi x, \qquad -1 < x < 1;$$

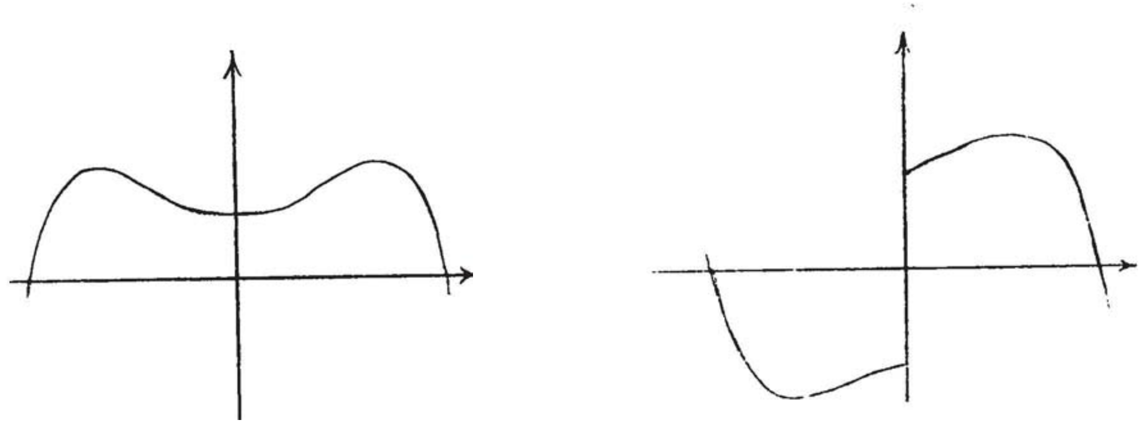
and hence that

$$\coth 1 = 1 + 2\sum_{1}^{\infty} \frac{1}{1+n^2\pi^2}.$$

4. Fourier Sine and Cosine Series

Let $f(x)$ be sectionally smooth on the interval $(0,L)$. In the last section we saw how we could represent $f(x)$ as a Fourier Series on this interval. In this section we shall see how $f(x)$ can be represented for x in the interval $0 < x < L$ by a different Fourier Series in which all the sine terms are absent or, alternatively, one in which all the cosine terms are absent. Before doing this we introduce some preliminary concepts.

Definition 4.1. A function $g(x)$ is said to be an even function if $g(-x) = g(x)$. (Examples: $g(x) = x^2$, $g(x) = x^4 + x^2$, $g(x) = \cos x$). A function $g(x)$ is said to be an odd function if $g(-x) = -g(x)$. (Examples: $g(x) = x^3$, $g(x) = x^3 + x$, $g(x) = \sin x$). See Figure 4.1.



An even function An odd function

Figure 4.1

Theorem 4.1. If $e_1(x)$ and $e_2(x)$ are even functions of x and $0_1(x)$ and $0_2(x)$ are odd functions then

$e_1(x)e_2(x)$ is an even function of x

$e_1(x)0_1(x)$ is an odd function of x

$0_1(x)0_2(x)$ is an even function of x ,

also

$$\int_{-a}^{a} e_1(x)dx = 2\int_0^a e_1(x)dx$$

and

$$\int_{-a}^{a} 0_1(x)dx = 0.$$

The proof is left to the student (Problem 4.1).

Suppose now that g(x) is sectionally smooth in the interval (-L,L). By formulas (3.5)-(3.7) we have

$$(4.1) \qquad \frac{g(x+) + g(x-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L},$$

where

$$(4.2) \qquad a_n = \frac{1}{L}\int_{-L}^{L} g(x) \cos \frac{n\pi x}{L}\, dx, \; b_n = \frac{1}{L}\int_{-L}^{L} g(x) \sin \frac{n\pi x}{L}\, dx.$$

From equation (4.2) we observe that if g(x) is odd then, by Theorem 4.1, $a_n = 0$ for all n; if g(x) is even then $b_n = 0$ for all n. To put it another way: An even function has only even terms (cosines) in its Fourier Series; an odd function has only odd terms (sines) in its Fourier Series.

Now suppose f(x) is given only on the interval $0 \leqq x \leqq L$. If we wish to represent f(x) by a <u>sine series</u>, we make the <u>odd periodic extension of f(x) from the interval (0,L)</u>. That is we define the function g(x) by

$$(4.3) \qquad g(x) = f(x), \; 0 \leqq x < L,$$

$$(4.4) \qquad g(x) = -f(-x), -L \leqq x < 0,$$

$$(4.5) \qquad g(x + 2L) = g(x), \text{ all } x.$$

(See Figure 4.2)

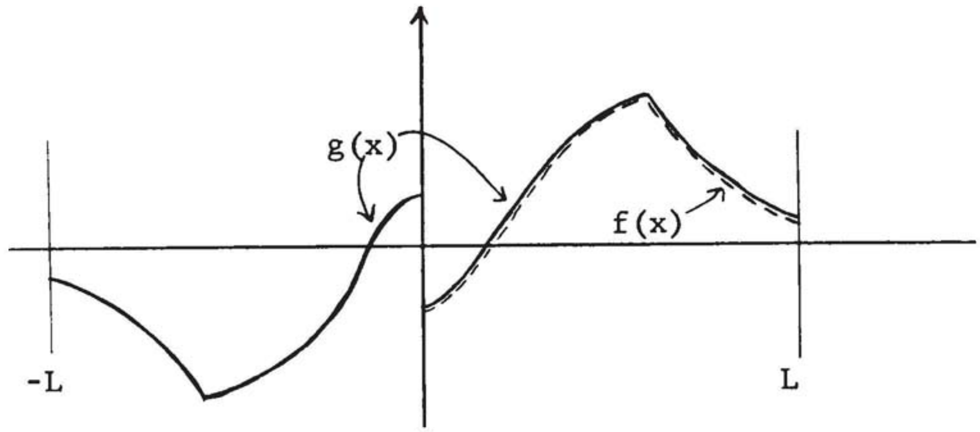


Figure 4.2

Then we expand g(x) in a Fourier Series on (-L,L) as in equation (4.1). Using Theorem 2.1 and equation (4.2) we get for $0 < x < L$,

$$\frac{f(x+) + f(x-)}{2} = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{L} ,$$

where

$$b_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi x}{L} \, dx.$$

Similarly if we wish to expand f(x) in a <u>cosine</u> <u>series</u> we obtain the result

$$\frac{f(x+) + f(x-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \operatorname{cox} \frac{n\pi x}{L} ,$$

where

$$a_n = \frac{2}{L} \int_0^L f(x) \cos \frac{n\pi x}{L} \, dx.$$

Problems

4.1 Prove Theorem 4.1.

4.2 Expand the function $f(x) = x, \quad 0 \leqq x < 1.$

(a) In a cosine series.

(b) In a sine series.

Sketch the first three terms of the sequence of partial sums in each case.

Answer: (a) $\frac{1}{2} - \frac{4}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \cos (2n+1)\pi x$,

(b) $\frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin n\pi x.$

4.3 Expand the function $f(x) = \sin x$, $0 \leqq x < \pi$ in

(a) a cosine series,

(b) a sine series.

4.4 Expand the function $f(x) = x^2$, $2 \leqq x < 3$ in

(a) a cosine series of period 2.

(b) a sine series of period 2.

Sketch the first three terms of the sequence of partial sums in each case.

4.5 (a) Find the formula for the coefficients in the Fourier sine series for the function $f(x)$ defined in Figure 4.3.

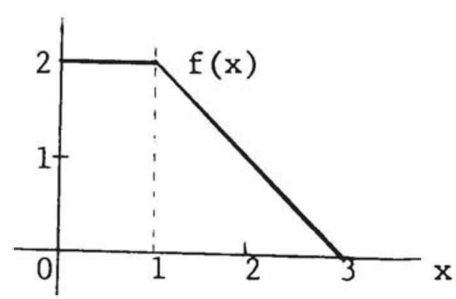


Figure 4.3

(b) To what must the series of part (a) sum when $x = 7/2$?

(c) It is desired to determine the <u>Fourier series for the function on the range</u> (0,3). Set up the integrals which must be evaluated to determine the numerical values of the Fourier coefficients a_n and b_n.

Partial Answers:

(a) $b_n = \frac{2}{n\pi}\left(2 + \frac{3}{n\pi}\sin\frac{n\pi}{3}\right)$.

(c) $$a_n = \frac{4}{3}\int_0^1 \cos\frac{2n\pi(x-3/2)}{3}\,dx + \frac{2}{3}\int_1^3 (3-x)\cos\frac{2n\pi(x-3/2)}{3}\,dx$$

4.6 Obtain the Fourier <u>Sine</u> Series for the function shown in Figure 4.4.

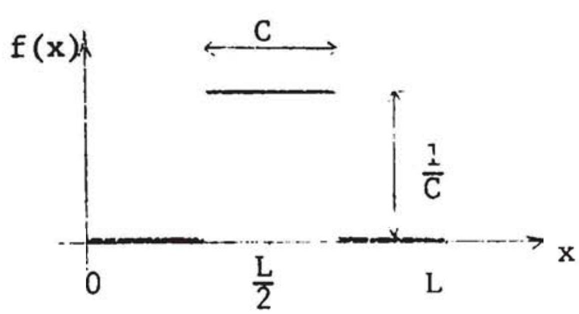


Figure 4.4

Answer: $\frac{4}{\pi C}\sum_{n\ \mathrm{odd}}(-1)^{(n-1)/2}\, n^{-1}\sin\frac{n\pi C}{2L}\sin\frac{n\pi x}{L}$.

4.7 (a) Obtain the Fourier Cosine Series Expansion for the function indicated in Figure 4.5.

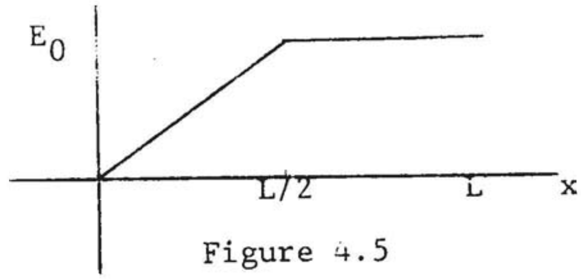


Figure 4.5

(b) To what number does the series in (a) converge when $x = 5L/4$?

4.8 (a) Find the Fourier Sine Series for the function

$$f(x) = \begin{cases} -1 \ , & 0 \leqq x < \frac{1}{2} \\ +1 \ , & \frac{1}{2} < x \leqq 1. \end{cases}$$

(b) Without summing the series, what does the series converge to when $x = 5/4$?

(c) Use the results of (a) and (b) to find the sum of the series

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} \ldots$$

Answer: (a) $-\frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\sin 2(2k+1)\pi x}{(2k+1)}$.

(b) -1. (c) $\pi/4$.

5. Applications to Solutions of Differential Equations

Fourier Series have many applications in differential equations. We have seen one such application in Section 1, and we illustrate another by an example. Consider the differential equation

(5.1) $$y'' - k^2 y = g(x),$$

where $g(x)$ is a periodic function of period 2π, and let us say is given by its Fourier Series

$$g(x) = \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} (\alpha_n \cos nx + \beta_n \sin nx).$$

We are looking for a periodic solution $y(x)$ of equation (5.1) with period 2π (the same as that of the "forcing function", $g(x)$). Let us proceed on the tentative assumption that this solution $y(x)$ can be represented by its Fourier Series:

$$y(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

and that this series can be differentiated termwise twice. Substituting in equation (5.1) we obtain

$$-\frac{k^2 a_0}{2} - \sum_{n=1}^{\infty}\left[(k^2+n^2)a_n \cos nx + (k^2+n^2)b_n \sin nx\right]$$

$$= \frac{\alpha_0}{2} + \sum_{n=1}^{\infty}(\alpha_n \cos nx + \beta_n \sin nx).$$

Equating the coefficients of like terms (or, alternatively, multiplying each side by cos mx or sin mx and integrating from $-\pi$ to π) we get

$$a_0 = \frac{\alpha_0}{k^2}$$

$$-(k^2+n^2)a_n = \alpha_n, \text{ or } a_n = \frac{-\alpha_n}{k^2+n^2} \qquad n \geqq 1$$

$$-(k^2+n^2)b_n = \beta_n, \text{ or } b_n = \frac{-\beta_n}{k^2+n^2} \qquad n \geqq 1.$$

Thus we are led to a solution of the form

$$(5.2) \qquad y(x) = \frac{-\alpha_0}{2k^2} - \sum_{n=1}^{\infty}\left(\frac{\alpha_n}{k^2+n^2}\cos nx + \frac{\beta_n}{k^2+n^2}\sin nx\right).$$

If α_n and β_n, the Fourier coefficients of the given function $g(x)$ go to zero like $\frac{1}{n^2}$, or faster, as $n \to \infty$ then the series in (5.2) can be differentiated termwise twice and thus one can conclude that (5.2) is actually a solution of equation (5.1). To be more specific suppose that $g(x)$ is the sawtooth function shown in Figure 5.1. Its Fourier Series is

$$\frac{\pi}{2} - \frac{4}{\pi}\left[\frac{\cos x}{1} + \frac{\cos 3x}{3^2} + \ldots + \frac{\cos(2n+1)x}{(2n+1)^2} + \ldots\right]$$

that is,

$$\alpha_0 = \pi$$

$$\alpha_n = 0 \quad \text{if } n \text{ is even,}$$

$$\alpha_{2n+1} = -\frac{4}{\pi(2n+1)^2}\,.$$

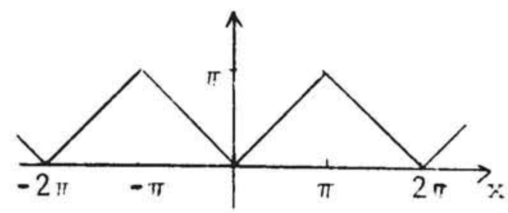


Figure 5.1

Consequently in this case equation (5.2) takes the form

$$(5.3) \qquad y(x) = \frac{-\pi}{2k^2} + \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{\cos(2n+1)x}{[k^2+(2n+1)^2](2n+1)^2}\,.$$

If the right side of (5.3) is differentiated termwise twice it becomes

$$(5.4) \qquad -\frac{4}{\pi}\sum_{n=0}^{\infty}\frac{\cos(2n+1)x}{k^2+(2n+1)^2}\,.$$

It can be shown that the series (5.4) represents $y''(x)$, where $y(x)$ is defined by equation (5.3). Subtracting k times the right hand side of equation (5.3) from the series (5.4) we get

$$y'' - k^2 y = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=0}^{\infty} \left[1 + \frac{k^2}{(2n+1)^2}\right] \frac{\cos(2n+1)x}{k^2+(2n+1)^2}$$

$$= \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\cos(2n+1)x}{(2n+1)^2} = g(x) \ ,$$

thereby verifying that equation (5.3) indeed gives a periodic function of period 2π which satisfies equation (5.1) when $g(x)$ is the sawtooth function.

Problems

5.1 With $g(x)$ the sawtooth function, as in the text above, find the Fourier Series of a periodic solution of period 2π of the equation

(a) $y''+k^2y = g(x)$. Show that resonance occurs if k is an odd integer.

Answer: $\dfrac{\pi}{2k^2} + \dfrac{4}{\pi} \displaystyle\sum_{n=1,3,\ldots} \frac{\cos nx}{n^2(n^2-k^2)}$.

(b) $y''+3y'+5y = g(x)$.

Answer: $\dfrac{\pi}{10} - \dfrac{4}{\pi} \displaystyle\sum_{n=0}^{\infty} \frac{-4(n^2+n-1)\cos(2n+1)x+3(2n+1)\sin(2n+1)x}{(2n+1)^2[25-(2n+1)^2+(2n+1)^4]}$.

5.2 A rectangular plate of dimensions 2×1, as indicated in Figure 5.2 has the edge $y = 0$ held at the temperature distribution $50 \sin \frac{\pi x}{2}$, and the edge $y = 1$ is held at $150 \sin \pi x$. The edges $x = 0$ and $x = 2$ are held at 0°C.

(a) Find the steady state temperature distribution in the place.

Answer:

$$\frac{150}{\sinh \pi} \sinh \pi y \sin \pi x$$

$$+ 50\left(\cosh \frac{\pi y}{2} - \coth \frac{\pi}{2} \sinh \frac{\pi y}{2}\right) \sin \frac{\pi x}{2} .$$

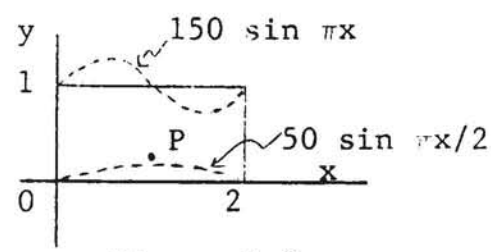


Figure 5.2

(b) Find the steady state temperature at the point $P(1, \frac{1}{4})$ to two significant figures.

Answer: 32°.

5.3 The block shown in Figure 5.3 has dimensions $7\times5\times3$ and is isotropic and homogeneous.

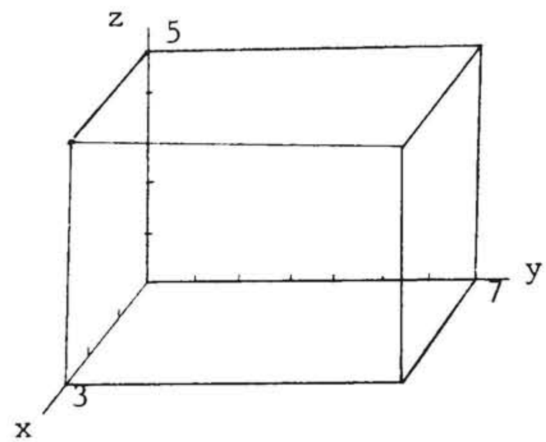


Figure 5.3

The five faces $z = 0$, $z = 5$, $x = 0$, $x = 3$, $y = 0$ are held at temperature zero, while the face $y = 7$ is kept at temperature $60 \sin \pi x \sin \pi z$. Use the method of separation to find the steady state temperature $u(x,y,z)$. Use this to find, to two decimal place accuracy, the steady state temperature at the point

$$\frac{3}{2}\ ,\ 7 - \frac{1}{\sqrt{2}\pi} \log 30\ ,\ \frac{3}{2}\ .$$ [Check: the answer turns out to be a nice even number.]

5.4 A metal rod, L centimeters long, with coefficient of thermal diffusivity α^2 is held in boiling water until it is heated uniformly to 100°C. Then (at time $t = 0$) it is withdrawn and the ends are pressed against ice so that the temperature of the ends is brought to 0°C and held at that temperature. Assume the lateral surface of the rod is insulated, so that there is no heat loss across the lateral surface and the problem can be treated as one-dimensional heat flow.

(a) Find the temperature $u(x,t)$ in the form of an infinite series.

Answer: $\frac{400}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} e^{-\left(\frac{(2k+1)\pi\alpha}{L}\right)^2 t} \sin \frac{(2k+1)\pi x}{L}$.

(b) If the coefficient of thermal conductivity is 1.8 cal/cm.sec.°C; the density is 10. grams/cm^3; the specific heat is 0.02 Cal/gram°C and the length L = 6 cm. find the temperature (to the nearest tenth of a degree) at the midpoint of the rod after 0.4 seconds.

Answer: 47.6°.

5.5 A simple harmonic oscillator of mass M and stiffness K is acted on by the pulsed periodic force F(t) shown in Figure 5.4. Determine the forced response of the oscillator (particular solution) to this excitation (in the form of an infinite series). From Problem 4.6,

$$F(t)= \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\sin \frac{n\pi}{2} \sin \frac{n\pi}{4}}{n} \sin \frac{n\pi t}{4}.$$

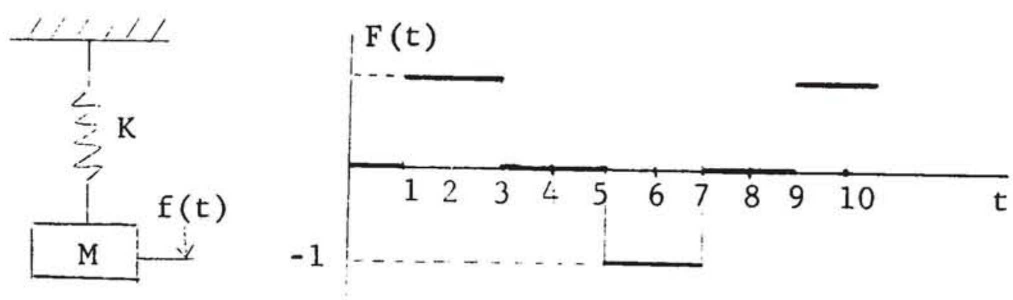


Figure 5.4

Answer: $\frac{2\sqrt{2}}{\pi M} \left[\frac{\sin(\pi t/4)}{\frac{K}{M} - \frac{\pi^2}{16}} - \frac{\sin(3\pi t/4)}{\frac{K}{M} - 9\frac{\pi^2}{16}} + \cdots \right]$, provided $\frac{4}{\pi}\sqrt{\frac{K}{M}}$ is not an odd integer. If it is an odd integer N, the corresponding term must be replaced by $\pm \frac{2t}{\pi N^2} \cos(N\pi t/4)$.

5.6 In the L-C circuit shown below the voltage $E(t)$ is the sawtooth function shown.

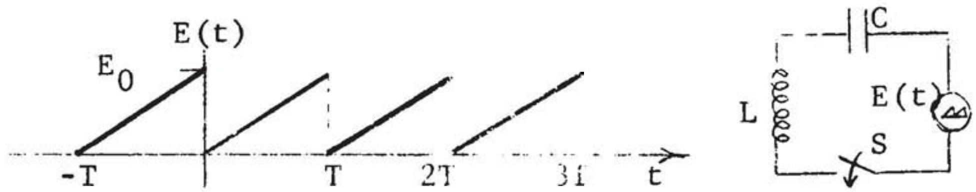


Figure 5.5

(a) Show that the Fourier series for this function is

$$E_0\left(1/2 - 1/\pi \sum_{n=1}^{\infty} 1/n \sin n\omega t\right), \text{ where } \omega = \frac{2\pi}{T}.$$

(b) Suppose that for $t \leq 0$ the switch S is open, so that at $t = 0$ there is no charge on the capacitor and no current flowing in the circuit (i.e., $q(0) = 0$, $\dot{q}(0) = 0$.). For $t > 0$ the switch is closed. Find $q(t)$ (in the form of an infinite series).

Answer:

$$E_0\left\{\frac{C}{2}(1-\cos\omega_0 t) + \frac{1}{L\pi\omega_0}\sum_{n=1}^{\infty}\frac{(n\omega\sin\omega_0 t - \omega_0\sin n\omega t)}{n(\omega_0^2-n^2\omega^2)}\right\}$$

where $\omega_0^2 = \frac{1}{LC}$.

5.7 (a) In Example 3.4 of Chapter 7 we derived the heat equation $\nabla^2 u = \frac{1}{\alpha^2}\frac{\partial u}{\partial t}$. Consider the heat flow in the rod shown in Figure 5.6. The curved surfaces are assumed

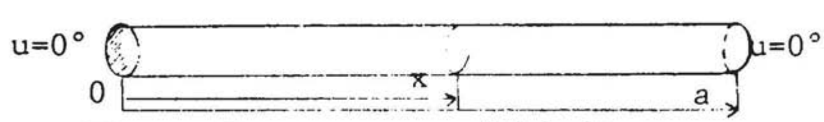


Figure 5.6

to be insulated and the temperature $u(x,t)$ a function of the distance x along the rod and the time t. Then $u(x,t)$ satisfies the equation

(5.5) $$\frac{\partial^2 u}{\partial x^2} = \frac{1}{\alpha^2}\frac{\partial u}{\partial t} .$$

Suppose that the temperature at the ends of the rod are fixed at 0°.

Then

(5.6) $$u(0,t) = 0 ,$$

(5.7) $$u(a,t) = 0.$$

If at time $t = 0$ the temperature is

(5.8) $$u(x,0) = f(x)$$

use the method in the text to find $u(x,t)$ satisfying (5.5)-(5.8).

[Hint: $\frac{X''}{X} = \frac{1}{\alpha^2}\frac{T'}{T} = -k^2$

$$X = A \sin\frac{n\pi x}{a} , \quad T = Be^{-\left(\frac{n\pi\alpha}{a}\right)^2 t}$$

$$u(x,t) = \sum C_n \sin\frac{n\pi x}{a} e^{-\left(\frac{n\pi\alpha}{a}\right)^2 t} .$$

To satisfy (5.8), $\sum C_n \sin\frac{n\pi x}{a} = f(x)$, from which

$C_n = \frac{2}{a}\int_0^a f(x) \sin\frac{n\pi x}{a}\, dx.$]

(b) If the initial temperature distribution is $x(a-x)$ deg. Centigrade and the rod is 50 cm. long and is made of copper (specific gravity = 8.89; thermal conductivity $k = .84$ cal/cm.deg.sec. and specific heat $c = .092$ cal/gr.deg.) find the temperature at the midpoint at $t = 10$ minutes. Answer: 57°.

(c) Same as (b) but the rod is made of glass ($\rho = 2.4$ gr/cm^3, $k = .0015$ cal/cm.deg.sec. and specific heat $c = .18$ cal/gr.deg.). Answer: 620°.

5.8 (a) In Problem 5.7 show that if the ends $x = 0$, $x = a$ are insulated then conditions (5.6), (5.7) are replaced by

$$\frac{\partial u}{\partial x}(0,t) = 0 \tag{5.9}$$

$$\frac{\partial u}{\partial x}(a,t) = 0. \tag{5.10}$$

(b) Find the temperature $u(x,t)$ in this case.
[Hint: Expand $f(x)$ in a Fourier Cosine Series.]
Answers: (a) 417°, (b) 613°.

5.9 A rod 2 feet long with coefficient of thermal diffusivity $\alpha^2 = 0.3$ sq. ft/hr. has its left end held at temperature 0° and its right end held at 80° until a steady temperature state is attained. Then (at time $t = 0$) the right end is suddenly brought to the same temperature as the left end, namely to 0°.

(a) Find the temperature distribution $u(x,t)$ in the rod.

Answer: $\dfrac{160}{\pi} \sum_{n=1}^{\infty} \dfrac{(-1)^{n+1}}{n} \sin \dfrac{n\pi x}{2}\, e^{-n^2\pi^2(0.075)t}$.

(b) Compute the temperature at the midpoint of the rod after twenty minutes to one decimal place accuracy. Answer: 38.3°.

5.10 The vibrations of a tightly sketched string may be analyzed as follows. Figure 5.7 shows a small section of the string and the forces acting on its ends (gravity is neglected in comparison with the tension T).

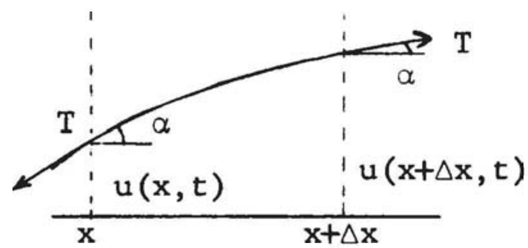


Figure 5.7

The net upward force is $(T \sin \alpha)_{x+\Delta x} - (T \sin \alpha)_x$. Equating this to the mass times the acceration $\rho \frac{(\Delta x)}{\cos \alpha} \frac{\partial^2 u}{\partial t^2}$ and letting $\Delta x \to 0$ show that

$$\cos \alpha \frac{d}{dx} (\sin \alpha) = \frac{\rho}{T} \frac{\partial^2 u}{\partial t^2} ,$$

where the tension is taken to be constant along the string.

Noting that $\tan \alpha = \frac{\partial u}{\partial x}$ show that

$$\frac{\partial^2 u}{\partial x^2} = \left(1 + \left(\frac{\partial u}{\partial x}\right)^2\right) \frac{\rho}{T} \frac{\partial^2 u}{\partial t^2} .$$

If we assume that $\frac{\partial u}{\partial x}$ is small, we can write

(5.11) $$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} .$$

This is called the <u>wave</u> <u>equation</u>.

(a) Show that $c = \sqrt{\frac{T}{\rho}}$ has the dimensions of velocity. It is the velocity of propagation of the waves of (5.11).

(b) If the string is fixed at its ends $x = 0$, $x = L$ and if at time $t = 0$ it is displaced to a position $u(x,0) = f(x)$ and released from rest then $u(x,t)$ satisfies (5.11) and the conditions

$$u(0,t) = 0$$

$$u(L,t) = 0$$

$$\frac{\partial u}{\partial t}(x,0) = 0$$

$$u(x,0) = f(x).$$

Use the method of separation of variables to find $u(x,t)$.

[Hint: $\frac{X''}{X} = \frac{1}{c^2}\frac{T''}{T} = -k^2$; $X = A \sin \frac{n\pi x}{L}$, $T = B \cos \frac{n\pi ct}{L}$

(5.12) $u(x,t) = \sum_{n=1}^{\infty} C_n \sin \frac{n\pi x}{L} \cos \frac{n\pi ct}{L}$;

where

$$C_n = \frac{2}{L}\int_0^L f(x) \sin \frac{n\pi x}{L}\, dx].$$

Interpret the individual terms in (5.12) as standing waves.

(c) Show that (5.12) can also be written in the form

(5.13) $u(x,t) = \sum_{n=1}^{\infty} \frac{C_n}{2} [\sin \frac{n\pi}{L}(x+ct) + \sin \frac{n\pi}{L}(x-ct)].$

Show that any function $f(x+ct)$ may be regarded as a wave traveling to the left with velocity v; any function $f(x-ct)$ as a wave travelling to the right with velocity v. Hence interpret the individual terms in the series of (5.13) as the superposition of waves.

5.11 (a) Show that the longitudinal vibration of a horizontal bar is described by the equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \text{ where } c = \sqrt{\frac{E}{\rho}} .$$

[Hint: In Figure 5.8 we examine a section of the rod of length Δx. Newton's law $F = ma$ becomes

$$(pA)|_x - (pA)|_{x+\Delta x} = \rho(\Delta x)A \frac{\partial^2 u}{\partial t^2},$$

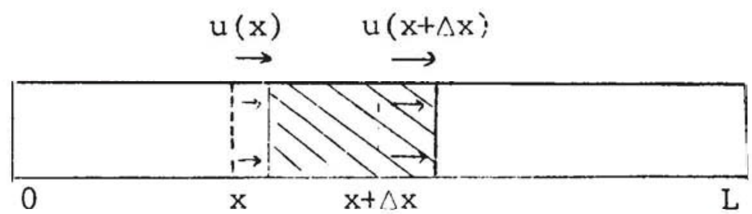


Figure 5.8

where p is the pressure, A the cross sectional area and ρ the density. Dividing by Δx and letting $\Delta x \to 0$ we obtain

$$(5.14) \qquad \frac{\partial(pA)}{\partial x} = (\rho A) \frac{\partial^2 u}{\partial t^2} .$$

Note from Figure 5.8 that

$$\text{strain} = \frac{u(x+\Delta x) - u(x)}{\Delta x} .$$

Hence by Hooke's law

$$(5.15) \qquad p = E \frac{\partial u}{\partial x},$$

where E is Young's modulus.

Use (5.15) in (5.14) and the assumption that the cross section is uniform to complete (a)].

What is the physical interpretation of the conditions:

(b) $\frac{\partial u}{\partial x}\Big|_{x=L} = \frac{S}{E}$.

(c) $u(0,t) = 0$.

(d) $u(x,0) = 0$.

(e) $\frac{\partial u}{\partial t}(x,0) = 0$.

5.12 In Problem 5.11 let $v(x,t) = u(x,t) - \frac{Sx}{L}$. Find the conditions on $v(x,t)$ corresponding to (a)-(e), and then solve for $v(x,t)$. Hence find $u(x,t)$, the solution to Problem 5.11.

5.13 If the rod in Problem 5.8 is carrying a current which is generating heat at a constant rate show that (5.5) is replaced by

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2} + r$$

where r is a constant. Show that a solution to this equation is $u(x,t) = v(x,t) + rt$ where $v(x,t)$ is a solution of (5.5).

5.14 If the rod in Problem 5.8 radiates heat by Newton's law show that (5.5) is replaced by

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2} - B(u-T_0) ,$$

where B is a positive constant. Show that this equation has a solution of the form $u(x,t) = e^{-Bt} v(x,t)$ where $v(x,t)$ satisfies (5.5).

5.15 The displacement $u(x,y,t)$ of a rectangular vibrating membrane satisfies the equation

$$\text{(i)} \qquad \frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right)$$

where $c^2 = \frac{Tg}{\rho}$ (ft/sec), T(lb/ft) is the tension, $g = 32$ ft/sec^2 and ρ(lb/ft^2) is the surface density.

(a) Derive (i) in the manner of Problem 5.10.

(b) Give the physical meaning of the following conditions

(ii) $u(0,y,t) = 0$

(iii) $u(a,y,t) = 0$

(iv) $u(x,0,t) = 0$

(v) $u(x,b,t) = 0$

(vi) $u(x,y,0) = f(x,y)$

(vii) $\frac{\partial u}{\partial t}(x,y,0) = 0$

(c) Use the method of separation of variables to find the formal solution of (i)-(vii):

$$u(x,y,t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} E_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \cos\left(\pi c \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2}}\, t\right)$$

where

$$E_{mn} = \frac{4}{ab} \int_0^b \int_0^a f(x,y) \sin \frac{m\pi x}{a} \sin \frac{m\pi x}{b}\, dx\, dy.$$

5.16 (a) Fill in the details of the following method of solving equation (5.1).

Expand $g(x)$ in the complex exponential form of the Fourier Series as in Problem 2.6,

$$g(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx}.$$

Find a particular solution of

$$y'' - k^2 y = e^{inx};$$

this solution will be

$$y = -\frac{1}{k^2+n^2} e^{inx}.$$

Using the Superposition Principle (assuming that it can be extended to infinite sums) equation (5.1) has a particular solution

$$y = \sum_{n=-\infty}^{\infty} -\frac{a_n}{k^2+n^2} e^{inx}.$$

This can be written in the form

$$y = -\frac{a_o}{k^2} - \sum_{n=1}^{\infty} \frac{1}{k^2+n^2} (a_n e^{inx} + a_n e^{-inx}).$$

If $g(x)$ is a real function then a_n and a_{-n} are conjugate complex numbers and

$$a_n e^{inx} + a_{-n} e^{-inx} = 2\mathrm{Re}(a_n e^{inx}),$$

so that finally

$$y = -\frac{a_o}{k^2} - \sum_{n=1}^{\infty} \frac{2}{k^2+n^2} \mathrm{Re}(a_n e^{inx}).$$

(b) Use this method to solve Problem 5.1.

5.17 A mass is connected with a spring and dash-pot as shown. The alternating force F(t) shown in Figure 5.10 is applied to the mass.

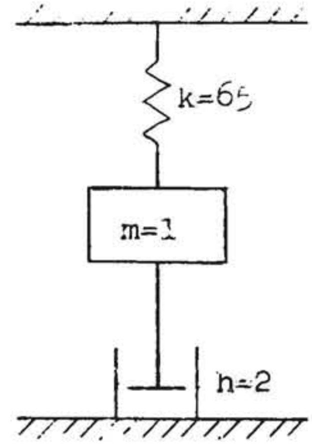


Figure 5.9

(a) Using the Fourier Series for F(t), (cf. Problem 2.2 (b) or 5.16) find the steady state response.

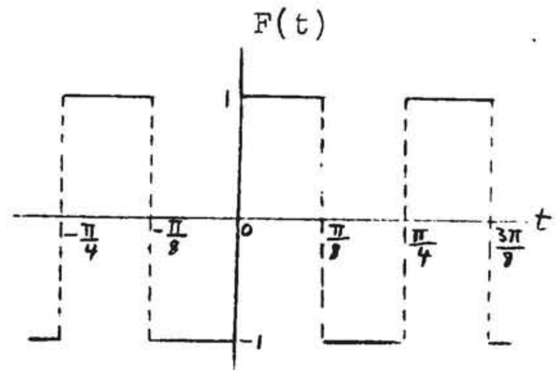


Figure 5.10

(b) Show that the amplitude of the steady state response is about $1/4\pi$ and is due almost entirely to the first term of the Fourier Series of F(t).

Index

Adjoint of a matrix, 5.113
an operator, 5.117
Analog, electro-mechanical, 6.3
Angle between two vectors, 5.11
Angular momentum, 7.14

Basis, orthogonal, 5.5, 5.16
orthonormal, 5.15
Blast loading, 6.49
Boundary value problem, 8.3

Canonical form, Jordan, 6.45
Cauchy's inequality, 5.14
Characteristic equation, 5.50, 6.6
multiple roots of, 6.145
Characteristic function, 5.50
Coefficients, Fourier, 8.21
Column vector, 5.33
Complementary solution, 6.48, 6.78, 6.154
Complex eigenvalues, 6.12
Component of vector orthogonal to subspace, 5.22
Condition, initial, 6.1
Configuration, relaxation, 6.33
Convergence Theorem for Fourier series, 8.25
Coordinate system, right-handed, 7.4
Cosine series, 8.40
Critical damping, 6.45
Cross product, 7.2
Curl of a vector field, 7.47
Cyclic permutation, 7.2

INDEX

Damping, critical, 6.45
Definite quadratic form, 5.99
∇, 7.30
∇^2, Δ, 7.51
Differential equation, 6.1, 8.46
    homogeneous, 6.1
    linear, 6.1
    non-homogeneous, 6.48, 6.76, 6.154
    n-th order, 6.141
    partial, 8.2
    second order, 6.101, 6.128
Direction cosines, ratios, 7.12
Directional derivative, 7.28
Divergence of vector field, 7.47
Dominant eigenvalue, 5.70
Dot product, 5.3

Eigenvalue, 5.46
    complex, 6.12
    of system of differential equations, 6.5
Eigenvector, 5.46
Equation, characteristic, 5.50
        multiple roots of, 6.145
        of system of differential equations, 6.6
    differential, see Differential equation
    heat, 7.53
    Laplace's, 7.81, 8.1
    Poisson's, 7.81
    wave, 8.56

Equipotential surface, 7.41
Even function, 8.41
Extension, periodic, of a function, 8.23

INDEX

Factor, integrating, 7.62
Field, scalar, 7.38
    vector, 7.39
Filter, 6.164
Fourier coefficients, 8.21
    cosine series, sine series, 8.40
    series, 8.17, 8.20
Frequency, natural, 6.9, 6.111
Function, characteristic, 5.50
    even, odd, 8.41
    orthogonal, 8.18
    periodic, 8.23
    ramp, saw-tooth, step, 8.34, 8.35

Gradient, 7.30

Half-rectified sine wave, 8.35
Harmonic oscillator, 6.1, 6.66, 8.54
Heat equation, 7.53
Hermitian form, matrix, 5.114
Hessian operator, 7.69
High-pass filter, 6.165
Homogeneous differential equation, 6.1

Identity, Jacobi, 7.16
Independent solutions of a differential equation, 6.24,6.114
Indefinite matrix, quadratic form, 5.99
Inequality, Cauchy's, 5.14
    Schwarz, 5.9
    triangle, 5.2
Inertia, 6.1, 6.3
Initial conditions, 6.1
Inner product, 5.3, 5.7

INDEX

Integrating factor, 7.62
Irrotational vector field, 7.55

Jacobi identity, 7.16
Jordan canonical form of a matrix, 6.45

Laplace's equation, 7.81, 8.1
Laplacian, 7.51
Left-hand limit, 8.23
Length of a vector, 5.2, 5.9
Level surface, 7.41
Limit, left-hand, right-hand, 8.23
Linear differential equation, 6.1
Low-pass filter, 6.1

Matrix, adjoint of, 5.113
    of system of linear differential equations, 6.4, 6.20
    Hermitian, 5.114
    orthogonal, 5.87
    symmetric, 5.58
    unitary, 5.114
Middle Factor, Rule of the, 7.11
Mode of oscillation, normal, 6.111
Mohole, 6.104
Moment of a force, of momentum, 7.14
Momentum, angular, 7.14
Multiple roots of characteristic equation, 6.145

Natural frequency, 6.9, 6.111
Negative definite matrix, quadratic form, 5.99
Norm of a vector, 5.9
Normal mode of oscillation, 6.111
Normal to a plane, 7.13

INDEX

Normalization of a vector, 5.11
Non-homogeneous differential equation, 6.48, 6.76, 6.154
n-th order differential equation, 6.141

Odd function, 8.41
Operator, adjoint, self-adjoint, unitary, 5.117
    Hessian, 7.69
Orthogonal basis, 5.15
    functions, 8.18
    matrix, 5.87
    projection, 5.22
    transformation, 5.87
    vectors, 5.11
Orthogonalization process of Schmidt, 5.16
Orthonormal basis, 5.6, 5.15
Oscillation, normal mode of, 6.111
Oscillator, harmonic, 6.1, 6.66, 8.54

Partial differential equation, 8.2
Particular solution, 6.48, 6.59, 6.154
Periodic extension of a function, 8.23
    function, 8.23
Permutation, cyclic, 7.2
Poisson's equation, 7.81
Positive definite matrix, quadratic form, 5.99
Potential, 7.55
Principle of superposition, 6.79
Product, cross, vector, 7.2
    inner, 5.3
    triple scalar, 7.7
Projection of a vector, 5.22

Quadratic form, 5.98

INDEX

Ramp function, 8.35
Rectified sine wave, 8.36
Relaxation time, 6.33
Resonance, 6.53, 6.58, 6.160
Right-handed coordinate system, 7.4
Rotation of a vector field, 7.47
Row vector, 5.33

Saw-tooth function, 8.34
Scalar field, 7.38
    product, 5.3
Schmidt orthogonalization process, 5.16
Schwarz inequality, 5.9
Second order differential equation, 6.101, 6.128
Sectionally smooth function, 8.25
Self-adjoint operator, 5.117
Semi-definite matrix, quadratic form, 5.99
Separation of variables, 8.12
Series, Fourier, 8.17, 8.20
    trigonometric, 8.17
Sine series, 8.40
    wave, rectified, half-rectified, 8.35
Solenoidal vector field, 7.51
Solution, complementary, 6.48, 6.78, 6.154
    particular, 6.48, 6.59, 6.154
Solutions of differential equation, 6.20
    independence of -, 6.24, 6.114
Square-cosine wave, 8.35
    -sine wave, 8.34
Step function, 8.35
Stiffness, 6.1, 6.3
Superposition principle, 6.79
Symmetric matrix, 5.58